

個体間相互作用に基づく創発的ダイナミクスに関する
エージェントベースモデリング

鈴木麗壘

名古屋大学 大学院人間情報学研究科

2002 年度

名古屋大学図書



11441253

目次

目次 I

図目次 IV

表目次 VI

第 1 章 序論 1

第 2 章 動的環境における進化と学習の相互作用 6

2.1 背景と目的 7

2.1.1 Baldwin 効果 7

2.1.2 Baldwin 効果に関する先駆的研究 8

2.1.3 動的環境における進化と学習の相互作用 11

2.1.4 目的 12

2.2 表現型可塑性を導入した繰り返し囚人のジレンマゲームにおける戦略の進化モデル
14

2.2.1 繰り返し囚人のジレンマゲーム 14

2.2.2 戦略の遺伝子表現 15

2.2.3 学習規則 15

2.2.4 繰り返し対戦と進化 17

2.3 基本的な進化実験 19

2.3.1 学習なしの戦略での実験 19

2.3.2 記憶長 2 ランダム型学習 20

2.3.3 記憶長 2 メタ・パブロフ学習 21

2.3.4 戦略の推移と Baldwin 効果 23

2.4 メタ・パブロフ[x00x]型戦略の解析 26

2.4.1 ESS 条件 26

2.4.2 状態遷移図に基づく解析 27

2.4.3 学習の役割 29

2.5 遺伝的記述の拡張 31

2.5.1 記憶長の突然変異を導入した進化 31

2.5.2 学習行列を遺伝子として取り込んだ進化 33

2.6 空間的局所性の影響 36

2.6.1 空間的局所性の導入 36

2.6.2 局所性と可塑性の進化 37

2.6.3 相互作用の規模と世代交代の規模 38

2.7 まとめ 42

第3章 協調行動の創発に対する空間的局所性の影響とその進化 45

3.1 背景と目的 47

3.1.1 空間的局所性と協調行動の創発 47

3.1.2 空間的局所性に関する2つの要素 48

3.1.3 空間的局所性の進化 49

3.1.4 N人版ゲームへの拡張 49

3.1.5 目的 52

3.2 空間的局所性を導入したN人版繰り返し囚人のジレンマゲームにおける戦略の進化モデル 54

3.2.1 N人版繰り返し囚人のジレンマゲーム 54

3.2.2 戦略の遺伝子表記 56

3.2.3 局所的な相互作用 57

3.2.4 局所的な世代交代 58

3.3 空間的局所性の基本的な影響 60

3.3.1 局所性の進化なしの場合での進化ダイナミクス 60

3.3.2 協調にかかるコストの影響 61

3.3.3 相互作用の規模の影響 62

3.3.4 世代交代の規模の影響 64

3.3.5 2つの局所性がもたらす多様性への影響 64

3.4 空間的局所性の進化 67

3.4.1 相互作用の規模の進化 67

3.4.2 世代交代の規模の進化 70

3.5 まとめ 72

第4章 行動多様性に対する情報共有の影響とその適応性 74

4.1 背景と目的 75

4.1.1 情報の多様性と行動の多様性 75

4.1.2	目的	76
4.2	イベント会場における混雑情報提供モデル	78
4.2.1	イベント会場	78
4.2.2	エージェントの行動	78
4.2.3	混雑情報の提供と情報共有	79
4.3	混雑に対する情報共有の影響	81
4.3.1	アトラクションの配置の影響	81
4.3.2	混雑情報共有の影響	83
4.4	情報共有の効果の変動と行動多様性	88
4.4.1	「突発待ち行列」の発生	88
4.4.2	情報共有による行動多様性維持の効果の減少	89
4.4.3	突発待ち行列の定量的解析	91
4.5	まとめ	93
第5章 結論 94		
謝辞 96		
参考文献 97		
関連発表論文 102		

目次

- 1.1 創発的ダイナミクス (模式図) 3
- 2.1 Baldwin 効果 7
- 2.2 Hinton と Nowlan の実験結果 9
- 2.3 メタ・パブプロフ学習の例 ($GS=[0001]$, $GP=[0011]$) 17
- 2.4 記憶長 2 学習なしでの実験結果 19
- 2.5 記憶長 2 ランダム学習での実験結果 20
- 2.6 記憶長 2 メタ・パブプロフ学習での実験結果 (2000 世代) 22
- 2.7 記憶長 2 メタ・パブプロフ学習での実験結果 (300 世代) 22
- 2.8 進化の過程における集団の可塑性と平均得点の相関 23
- 2.9 集団の可塑性・平均得点と進化の過程で出現した戦略の分布 24
- 2.10 メタ・パブプロフ[x00x]型戦略と他の記憶長 2 の戦略との対戦成績 27
- 2.11 メタ・パブプロフ[x00x]型戦略の状態遷移図 28
- 2.12 記憶長の突然変異を導入した場合の実験結果 31
- 2.13 学習行列を遺伝子に取り込んだ場合の実験結果
(メタ・パブプロフ[x00x]型戦略に収束した例) 34
- 2.14 学習行列を遺伝子に取り込んだ場合の実験結果
([x001], $GT=[1000]$ に収束した例) 34
- 2.15 4000 世代目において各戦略が集団中を占めた割合 (60 試行) 35
- 2.16 pPAVLOV (左) と[x001], $GT=[1000]$ (右) の状態遷移図 35
- 2.17 相互作用の規模 i によって定義される近傍 36
- 2.18 集団の可塑性の推移 37
- 2.19 平均得点の推移 37
- 2.20 集団の可塑性の最高値 39
- 2.21 平均得点の最高値 39
- 2.22 局所性と戦略の分布 40
- 2.23 2 つの局所性の変化に対する可塑性の最高値の比較 41
- 2.24 進化と学習の 2 つの適応プロセスが形作るイメージ 43
- 3.1 空間的局所性の導入 47
- 3.2 N 人版ゲームの利得式 54
- 3.3 N 人版ゲームの具体的状況設定の例 55
- 3.4 N 人版ゲームでの戦略の例 56
- 3.5 集団の構成 57
- 3.6 相互作用に関するグループ 58
- 3.7 世代交代に関するグループ 59

- 3.8 $I=8, R=8$ での進化の過程 60
- 3.9 協調者の割合に対する I, R の影響 61
- 3.10 協調者の割合の増加に対するコスト k の影響 62
- 3.11 $I=2, R=8$ の場合での協調者の割合の分布 63
- 3.12 $I=32, R=8$ の場合での協調者の割合の分布 63
- 3.13 $I=8, R=2$ の場合での協調者の割合の分布 66
- 3.14 $I=8, R=32$ の場合での協調者の割合の分布 66
- 3.15 相互作用の規模 i を進化, $R=8$ の場合での進化の過程 67
- 3.16 協調者の割合とグループの大きさの相関 68
- 3.17 相互作用の規模の進化を導入したときの協調者の割合 69
- 3.18 世代交代の規模 R を進化, $I=8, \beta=0.4$ の場合での進化の過程 70
- 3.19 協調者の割合と世代交代の規模の進化に対する β の影響 71
- 4.1 シミュレーション実行画面 81
- 4.2 エージェントのモードの推移 ($r=0.0$) 82
- 4.3 各アトラクションの待機エージェント数の推移 ($r=0.0$) 82
- 4.4 各アトラクションの待機エージェント数の推移 ($r=0.1$) 85
- 4.5 各アトラクションの待機エージェント数の推移 ($r=0.2$) 85
- 4.6 各アトラクションの待機エージェント数の推移 ($r=0.3$) 85
- 4.7 各アトラクションの待機エージェント数の推移 ($r=0.5$) 86
- 4.8 各アトラクションの待機エージェント数の推移 ($r=0.8$) 86
- 4.9 各アトラクションの待機エージェント数の推移 ($r=1.0$) 86
- 4.10 混雑情報提供確率 r と待機エージェント数の標準偏差 87
- 4.11 各アトラクションの待機エージェント数の推移 ($r=0.5, N_p=20000$) 88
- 4.12 各アトラクションの観賞予定エージェント数の推移 ($r=0.5, N_p=20000$) 89
- 4.13 10000 ステップ目で操作を加えた場合の待機エージェント数の推移
($r=0.5, N_p=20000$) 90
- 4.14 10000 ステップ目で操作を加えた場合の観賞予定エージェント数の推移
($r=0.5, N_p=20000$) 91
- 4.15 混雑情報提供確率 r と待機エージェント数標準偏差,
および, 突発待ち行列発生回数 (30 試行, $r=0.5, N_p=20000$) 92

表目次

- 2.1 囚人のジレンマゲームの利得行列 14
- 2.2 メタ・パブプロフ学習行列 16
- 2.3 全面裏切り戦略と[x00x]との対戦例 29
- 2.4 [x001]と[x00x]の対戦例（下線はノイズ） 29
- 2.5 [x00x]同士の対戦での協調の回復例（下線はノイズ） 30
- 2.6 学習行列を表す遺伝子列 GT の定義 33
- 3.1 囚人のジレンマゲームの利得行列 50
- 3.2 相互作用の規模の進化を導入した場合の平均値 69

第1章 序論

現実世界には、予測がつかないような複雑な現象が数多く存在する。簡単な例として、えさを運ぶアリの行列を考えてみる。アリは、えさを見つけると、自分たちの巣からえさまできれいに行列を作り、大勢でたくさんのえさを巣まで運ぶ。この行列を指揮しているような特別なアリは明らかに存在せず、個々のアリは勝手に振舞っているにもかかわらず、彼らはどのようにして行列というある種の秩序立った振る舞いを実現しているのだろうか。また、どのような進化の道りを経て、このような振る舞いを身に着けたのだろうか？このように、要素を見ただけでは全体の挙動が予測できないような系は、現実世界には無数に存在する。単純なものの集まりを全体として予測できない複雑なものにしている要因は一体何であろうか？また、そのような系は、どのような過程を経て生じるのであろうか？これが、本研究の根本的な問いである。

世の中に存在する複雑な現象については、複雑系科学と総称される様々な研究領域において議論されている。その内でも、中心的な論点の一つとなっているのが、上記のような創発的な振る舞いに関する議論である[有田 02]。創発という言葉は、「それ以前に見られなかった（予想できなかった）特性の出現」という一般的な意味づけのみならず、「多数の要素から構成される系において、要素間の相互作用によって、要素のみからでは説明できない大域的な挙動が表出すること、また、その大域的な挙動が更に要素間の相互作用に影響を及ぼすこと」、といった哲学的な背景に基づく意味づけや、「それ以前は存在しなかった構造（定常化した回路や反応なども含む）が時間経過に従って出現する」といった生物学的（進化論的）な意味づけがなされており、それぞれの文脈で以前から用いられてきた。

上述のアリの行列の例は、実際には次のような創発的なメカニズムによってもたらされている。アリは、えさを見つけると、フェロモンを出しながら巣に持ち帰る。このフェロモンはほかのアリを誘い出す働きがあるため、その道にはアリがさらに集まる。そのアリもえさを見つけるため、同時にフェロモンも増えていく。これを繰り返すうちに、えさと巣を結ぶ最短距離の道が出来上がる。つまり、フェロモンを介したアリ同士の間接的な相互作用によって、行列という大域的な挙動が創発したとすることができる。しかし、このようなアリの集団の振る舞いは、アリ一匹を見ただけでは全く予測がつかない。

アリの行列は、系に大域的な挙動を生み出すという意味での創発現象の典型的な例の一つである。また、例えば、進化の過程で様々な種が作り出されたり、新生児の知能が育ち、成長して成人として振る舞うようになるという過程は、新たな構造が出現するという意味での創発現象を端的に表すものとも言える。現実世界には、生命、自然、社会など、大小を問わずほぼありとあらゆる系の中に、上記のような個体（要素）間相互作用に基づく創発性が少なからず含まれていると考えられる。また、それらを

駆動する進化・学習に代表されるような適応メカニズムにも、新たな構造を生み出す創発性が含まれると考えられる（当然、生み出される構造が創発的であることもありうる）。したがって、それらの振る舞いを明らかにするには、先に挙げたような創発現象の統合的な解明が必要不可欠であると言える。

このような創発現象は、以前から事実として確認されてきたものの、単純なルールに従う要素が数多く集まることで複雑で予測不可能なパターンが生じるという分散的なものの見方は、西洋科学の基本的なパラダイムである要素還元主義的な見方とは大きく異なっている。したがって、この見方に基づいて議論を行うには、これまでとは異なる新しい方法論が必要であった。それが、構成論的手法と呼ばれる、ミクロなルールの記述を行ったモデルを主に計算機上に実装し、それを動かすことによって創発現象を実際に起こし、観察することで理解を深めるというアプローチである。近年、構成論的手法に基づくアプローチは、計算機の処理性能の飛躍的向上と、遺伝的アルゴリズムに代表されるような生物の進化や学習などの適応メカニズムの定式化が成熟したことによりはじめて創発現象を理解するための強力なツールとなった。人工生命研究は、創発性を生命の生命らしさの主要な鍵であると捉え、本質を捉えた抽象モデルを構築することによって、生命現象に関わる創発的な振る舞いについて多くの普遍的な知見を得ている。

本研究は、創発現象が生じるメカニズム、つまり、創発的ダイナミクスに関して、エージェントベースモデリングと呼ばれる構成論的アプローチを用いて、その普遍的な性質を明らかにすることを目的とする。エージェントベースモデリングとは、社会や生態系、生物集団といった多数の自律的な主体によって構成される集団を、主体間の相互作用のルールによってモデル化する手法である。本研究では、個体間相互作用に基づく創発的ダイナミクスに関わる次の3つの本質的な問いに対し、それぞれについてエージェントベースモデリングに基づく計算機実験を用いた具体的研究事例を通して知見を得ることを目的とする（図：1.1）。

- 1) 一つの系において、それを駆動する適応プロセスが複数の種類存在するとき、それらはどのように影響しあうのか？また、適応プロセス間の相互作用は系全体の挙動にどう反映されるか？
- 2) 個体間で行われる相互作用の構造やその強さ、例えば、ある個体が他の個体に直接影響することのできる範囲といったものが、系全体にどのような挙動をもたらすか？また、その構造自体の変化が系に及ぼす影響とはどのようなものか？
- 3) 集団において、個体間で共有する情報とそれを元に実際に行われる行動といった異なるレベルにおいて生じる多様性は、どのように影響し合っているか？

まず、1つ目の問いへの具体的研究事例として、進化と学習という2つの適応プロ

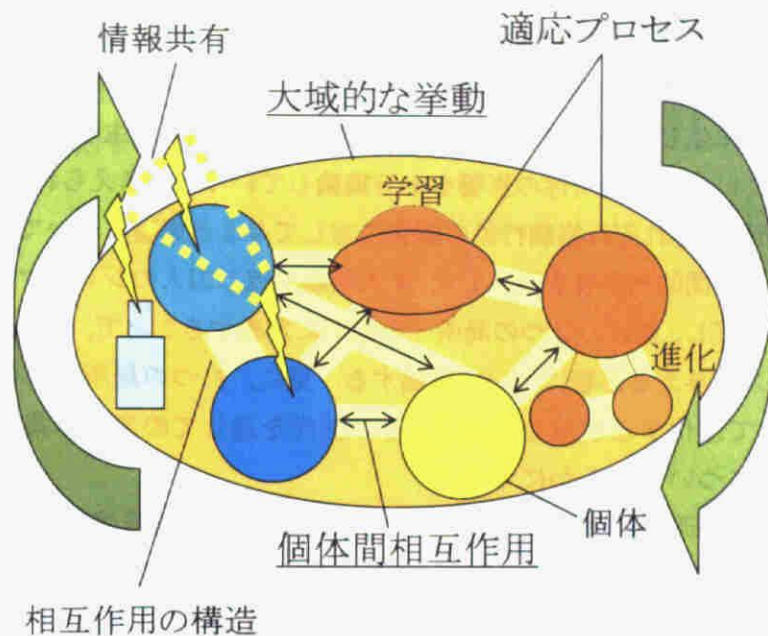


図 1.1: 創発的ダイナミクス(模式図)

セスの相互作用によって生じる創発的ダイナミクスを取り上げる。進化と学習は生物における代表的な適応プロセスである。その両者がいかに影響しあっているかという問題は 100 年以上前に端を発する生物学的問題であり、近年、この問題は進化論的計算、あるいは人工生命の研究領域における主要テーマの一つになってきた。中でも、Baldwin 効果と呼ばれる現象は、Hinton と Nowlan による先駆的な進化的計算機実験以来、生物学的知見のみならず工学的応用への展開についても議論されており、注目を集めている。本研究では、これまであまり議論されることがなかった、動的環境における進化と学習の相互作用に関し、表現型可塑性進化を導入した繰り返し囚人のジレンマゲームにおける戦略の進化モデルを用いることで、その特性を明らかにすることを目的とする。具体的には、自身の振る舞いが集団中の他の個体に大きく影響するゲーム論的状況設定において、個体の学習、すなわち表現型可塑性の進化が、協調行動の創発にどのように寄与するかについて、Baldwin 効果を中心に詳細に論ずる。また、ゲーム理論的な観点から、出現した個体の特徴について詳細に解析を行い、進化の結果得られた学習能力が協調行動を促進するメカニズムを明らかにする。更に、戦略の遺伝的記述の拡張や学習行列の進化、局所性の導入など、モデルをより一般化した条件での進化と学習の相互作用について詳細に論ずる。

2 つ目の問いに関しては、社会的集団において空間的な局所性とその進化が協調行動の創発に与える影響に焦点を当てる。これまで、社会的集団における協調行動の創発の起源やそのメカニズムの解明をめぐり、多くの繰り返し囚人のジレンマゲームにおける戦略の進化に関する研究が行われてきた。その中で、社会的・生態学的に妥当な状況設定での議論の展開を目指して、個体間の相互作用に空間的局所性を導入した議論

がなされている。これは、現実世界では必ずしも個体間の相互作用の関係が均一ではなく、ある程度の局所性を持って相互作用が行われているためである。しかし、従来の研究は、局所性に関して区別して考慮することのできる2つの要素、すなわち個体がゲームを行う対戦相手を選択する際に生じる局所性と、次世代の個体を選択する際に生じる局所性を同一のものとして扱っている点で局所性の影響を十分議論していないと考えられる。本研究は、これら2つの局所性がそれぞれ協調行動の創発に対して与える影響について知見を得ることを目的として、空間的局所性を導入したN人版繰り返し囚人のジレンマゲームの戦略の進化モデルを構築し、前述の2つの局所性を個別に定義することで、その違いが戦略の進化ダイナミクスに与える影響について議論する。更に、2つの局所性が各個体の持つ戦略の一部として進化する状況を設定し、その世代を通しての変化が集団全体の挙動にもたらす影響について明らかにする。

3つ目の問いに関する研究事例として、イベント会場における混雑情報提供を取り上げる。大規模テーマパークなどのイベント会場において、携帯情報端末などを用いて来場者に混雑情報を提供し、混雑を緩和する試みが検討されている。しかし、例えば、ある時点で「あるアトラクションが空いている」という情報を過剰な数の来場者が共有し、それに従って行動した場合、結果としてそのアトラクションが混雑してしまうことがありうる。これは、集団においてある情報を共有することが、必ずしも系全体にとって良い結果をもたらすとは限らないことを示す一例であり、系に生じる多様性の適応性やその起源・維持のメカニズムの解明は、複雑系科学における主要な研究テーマの一つである。本研究では、群集行動に対する情報共有の影響、特に、系全体において共有される情報の多様性と行動の多様性の間に生じる相互作用に注目し、両者の関係がもたらす創発的ダイナミクスを明らかにすることを目的として、イベント会場における混雑解消のための混雑情報提供を集団における行動の多様性に対する情報共有の影響とその適応性に関する具体的な状況設定としてモデル化し、マルチエージェント・シミュレーションを行う。具体的には、会場に設置された複数の施設を多数の来場者が各自の嗜好に基づいて観賞してまわる抽象モデルを構築し、混雑情報の提供による来場者間の情報共有が会場内の混雑の偏りに与える影響を明らかにする。更に、情報共有の行動多様性に対する効果が、時間経過に伴って大きく変動する現象について示し、その要因について詳細に論ずる。

以上のように、各テーマはそれぞれ具体的な状況設定における議論であるものの、それぞれ、先に挙げたような創発的ダイナミクスに関する本質的な問いに対して、重要な知見をもたらすことが期待される。

本論文の次章以降の構成は次のとおりである。2章では、進化と学習の相互作用について、表現型可塑性を導入した繰り返し囚人のジレンマゲームにおける戦略の進化実験をもとに議論する。3章では、個体間の局所的な相互作用について、空間的局

所性を導入した N人版繰り返し囚人のジレンマゲームの戦略の進化モデルを用いて明らかにする。4章では、集団内で共有される情報と行動の多様性の関係について、イベント会場における混雑解消のための混雑情報提供モデルを用いて議論する。5章で、議論を総括する。

第2章 動的環境における進化と学習の相互作用

進化と学習は生物における代表的な適応プロセスである。その両者がいかに影響しあっているかという問題は 100 年以上前に端を発する生物学的問題であり、これに関する重要なトピックに、Baldwin 効果と呼ばれる現象がある[Baldwin 1896]。この現象は、Lamarck 的な獲得形質の遺伝の仕組みが無くても、集団における個体の学習が集団全体の進化に方向性を与え、その結果学習によって獲得されていた形質が次第に生得的な形質へと進化していくというものである。

これまで、Baldwin 効果に関する進化実験による研究は、その存在を明確にした Hinton と Nowlan による先駆的な進化実験[Hinton 87]をはじめとして最適解が固定されたものがほとんどであり、動的な環境において Baldwin 効果がどのように働くかについての議論は不十分であった。しかし、現実世界においては、学習はむしろ動的な環境において有効に働くと考えられるため、動的な環境における進化と学習の相互作用を明らかにすることは重要であると言える。

そこで本章では、特に個体間の相互作用に適応度が依存した動的な環境として繰り返し囚人のジレンマゲームの戦略の進化を取り上げ、戦略に提案するメタ・パブプロフ学習に基づく表現型の可塑性を導入したマルチエージェントモデルを構築し、進化の過程を解析することで、動的な環境における進化と学習の相互作用に関する創発的ダイナミクスについて知見を得ることを目的とする。はじめに、基本的な進化実験において、このような動的な環境においても Baldwin 効果が有効に働き、集団は協調的で安定な戦略集団へ進化したことを示す。次に、戦略の推移と Baldwin 効果の関係、および進化の過程で出現し最終的に集団中を占めた安定な戦略であるメタ・パブプロフ [x00x]型戦略について解析する。更に、モデルをより一般化したオープンエンドな進化実験、局所的な相互作用を導入した実験結果について示し、考察する。

本章の構成は次のとおりである。2.1 節では、進化と学習の相互作用の一つである Baldwin 効果と関連する研究を紹介したうえで、本研究の目的について述べる。2.2 節では構築したモデルについて詳細に解説する。2.3 節では基本的な進化実験の結果および戦略の推移の解析について示し、メタ・パブプロフ[x00x]型戦略についての解析を 2.4 節で行う。2.5、2.6 節では、遺伝的記述の拡張や局所的な相互作用といった、モデルをより一般化した進化実験の結果について示す。最後に 2.7 節において、これまでの結果および考察をまとめるとともに、今後の課題について述べる。

2.1 背景と目的

2.1.1 Baldwin 効果

進化と学習の相互作用について、様々な議論がなされてきた。そのうち最も大きなトピックのひとつに、Lamarck 的な獲得形質が遺伝する仕組みの存在の是非をめぐる議論があった。現在では、そのようなメカニズムの存在は一般的に否定されているが、獲得形質の遺伝の仕組みが無くても、自然選択のみでそれと類似した効果が得られるという進化のシナリオが Baldwin によって提案された [Baldwin 1896].

(Baldwin 自身はこれを Organic Selection と呼んだが、その後 Baldwin 効果という名称で定着した。) これは、進化と学習が相互に与える影響を、学習のメリットとコストのバランスから説明するものであり、現在の一般的な定義では、次の 2 つの段階を経て、学習により獲得されていた形質が次第に生得的な形質へと進化していくものとされている [Turney 96].

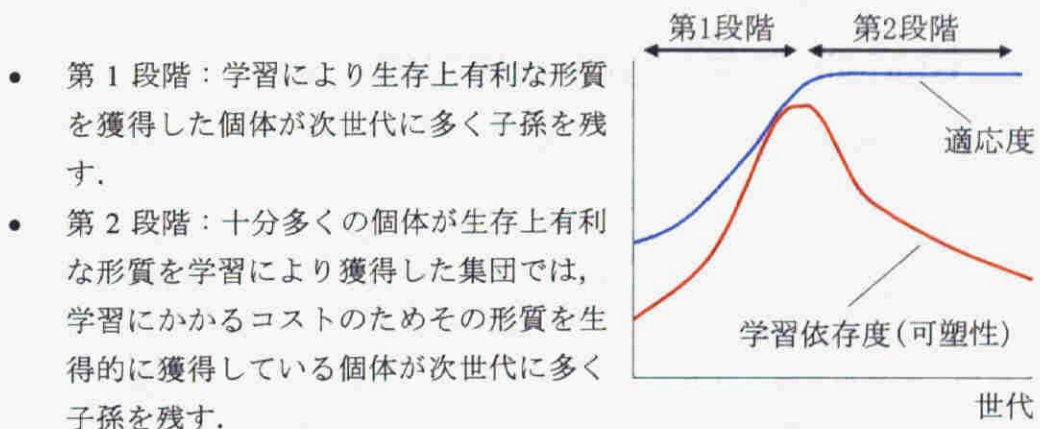


図 2.1: Baldwin 効果

なお、ここで述べる学習とは、人間の知能に代表されるような高度な学習メカニズムだけでなく、たとえば運動による筋肉の増強や日焼けによる皮膚の色の変化などを含む表現型の生涯の変化、すなわち広い意味での表現型可塑性 [Pigliucci 01] を示すものとする。

第 1 段階は、学習によって適応的な形質を獲得するメリットが中心的な選択圧として働き、学習可能な個体が集団中に広まっていく過程である。一見すると、第 1 段階により全ての個体が学習可能になれば適応度差は生じることはなく、それ以上進化は進まないように思われる。しかし、Baldwin 効果の興味深い点は、学習にかかるコストによる適応度の違いに注目した第 2 段階が存在することである。

多くの場合、学習には何らかの形でコストがかかる。例えば、形質を獲得するの

に時間やエネルギーがかかってしまうことや、そもそも学習し損ねてしまうといったことなどが考えられる。このコストは学習によるメリットと比べてわずかかもしれないが、確実に適応度の差に現れ、集団はコストが小さくなる方向へ徐々に進化すると考えられる。コストが最も小さい状況は、究極的には学習すべき形質を生得的に獲得している状況であるため、2つの段階を経て、はじめは学習によって獲得されていた形質が生得的に獲得されている形質へと進化するシナリオを描くことができる。

このとき、2つの段階において、集団全体の学習に対する依存度（可塑性）というものが定義できるとすれば、典型的には図 2.1 のような適応度と依存度のカーブを描くと考えられる。また、もしその形質についての学習能力の獲得と形質自体の生得的な獲得に遺伝的な相関がある、あるいは、遺伝子型空間と表現型空間内の距離関係に相関があるような場合、2つの段階はその形質の生得的獲得を促進する可能性があるとして起る現象は、Waddington が提案する遺伝的同化作用に相当する[Waddington 42]。Baldwin によってこの効果が提案されたのは100年以上前のことであるが、その存在を確認するのが難しいために、これまであまり注目されてこなかった。

2.1.2 Baldwin 効果に関する先駆的研究

しかし近年、Hinton と Nowlan による先駆的な進化実験によって、明確にされて以来、この効果は生物学的側面からだけでなく、進化的計算などの工学的分野からも注目されるようになり、新たな局面を迎えている。Hinton と Nowlan は Baldwin 効果を確認するために、遺伝的アルゴリズムを用いた抽象モデルを用いて進化実験を行い、特に Baldwin 効果の第1段階について明快に示した[Hinton 87]。

彼らが行った進化実験とは次のようなものである。集団における各個体は“1”，“0”，“?”の3種類の遺伝子が存在する20個の遺伝子列を持つ。初期集団（1000個体）における各個体の遺伝子は1:0:? = 1:1:2の比率でランダムに決定される。遺伝子列は各個体の適応度を決定するネットワークの結合のパターンを表す。“1”は正しい結合，“0”は誤りの結合であり、遺伝子が0または1であることは、結合がそれぞれ遺伝的に0または1に決定されていることを示す。遺伝子が“?”であることは結合が遺伝的には未決定であることを示し、この部分における正しい結合を学習プロセスによって後天的に獲得する。学習プロセスおよび適応度の計算方法は次のとおりである。

- A) 遺伝子中の“?”すべてについて、“1”または“0”をランダムに当てはめる。
- B) 正しいネットワークの結合（この場合はすべて“1”）と比較し、すべての結合が正しい結合と一致すれば、学習が完了したのものとしてCへ移る。そうでなければAに戻る。このとき*i*を学習が完了するまでに行った試行数とする。

ただし最大試行回数の 1000 回まで試行しても解と一致しない場合は $i=1000$, “?” が存在せず初めから解と一致している場合は $i=0$ とする.

C) 以下の式を用いて適応度を計算する.

$$\text{適応度} = 1 + \frac{19(1000 - i)}{1000} \quad (2.1)$$

(2.1) 式から, すべての結合が遺伝的に正しい個体の適応度は最大値 20, 遺伝的に誤りを含む個体の適応度は最小値 1, 正しい結合を学習によって獲得する可能性のある個体の適応度はこの間の値をとることになる. 選択は, 適応度に応じたルーレット選択で親 2 個体を選び, 一点交叉してできた子孫のうち片方を次の世代の個体とするものとする. これを 1000 回繰り返す, 次世代の集団を生成する.

彼らは, 以上のようなモデルで実験を行い, 図 2.2 に模式的に示すような進化が観察されたと報告している. 彼らは, 全遺伝子中に占める “0”, “1”, “?” の割合の推移に注目した. 初期数世代まではすべての個体が学習に失敗し, 選択圧がかからない状態が続いた. その後, 学習によって正しい結合の獲得に成功した個体が出現し, 集団中に広がり始めた. このとき, 全遺伝子中に含まれる 0 の割合は減少し, “1” と “?” の割合が増加した. これが学習によるメリットが働いた状態で, Baldwin 効果の第 1 段階である. その後, “?” の割合が若干減少したところで, 進化は収束した. この “?” の若干の減少と “1” の割合の増加が Baldwin 効果の第 2 段階である.

彼らの実験において注目すべき点は, 学習の結果が次世代に遺伝する仕組みが無いにもかかわらず, 学習が集団の進化に大きな影響を与えている点である. この結果では最終的に遺伝的に正しい結合の割合が約 60% もの集団へと進化したが, 仮に, この実験において学習を導入しなかった場合, すなわち遺伝子の値として “0” または “1” だけを用いる設定で実験を行った場合, 唯一高い適応度を持ったすべての遺

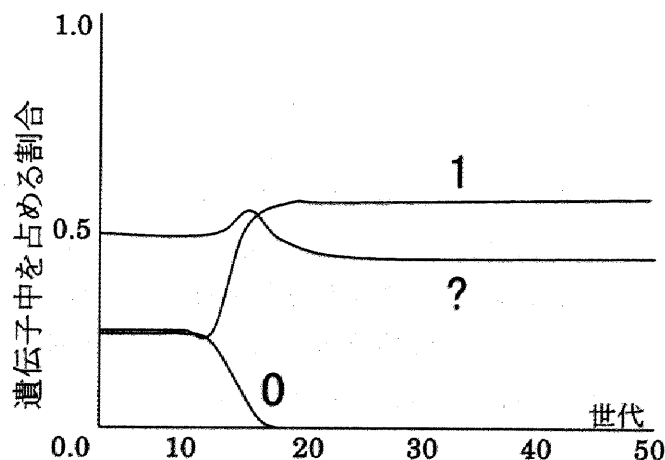


図 2.2: Hinton と Nowlan の実験結果

伝子が“1”である個体を $(1/2)^{20}$ の確率で発見しなくてはならないため、選択圧がかからないままの状態が続くことは容易に想像できる。この結果から、彼らは、学習結果が次世代に伝わらなくても最適解への進化を促進するこのような働きについて、尖った適応度地形をなだらかにする役割があると指摘している。なお、この実験では Baldwin 効果の第 2 段階である、“1”の増加と“?”の減少が大きく現れていないこともわかる。その理由は、彼らのモデルでは突然変異を用いていないために、第 2 段階が十分働く前に集団が 1 種類の遺伝子構成に収束し、進化が収束したことによると推測される。実際、Harvey や有田の行った突然変異を導入した追実験では、第 2 段階は継続され、徐々に正しい結合を多く持った集団へと進化していった[Harvey 93], [有田 02]。これらの実験結果は、Baldwin 効果の存在を明確に示したのと同時に、進化と学習が組み合わさったメカニズムの工学的応用の可能性を示したものである。

Baldwin 効果に関するもう一つの代表的研究として、Ackley らの進化的強化学習に関する研究を挙げる。彼らは、2次元空間に作られた生態系の中で、エージェントが敵を回避しつつ食べ物を手に入れるタスクについて、行動に対する先天的な評価基準を用いて学習する強化学習法 (Evolutionary Reinforcement Learning) を採用し、集団を進化させる実験を行った[Ackley 91]。各エージェントは外部からの視覚情報に対してとるべき行動を出力するニューラルネットワークと、入力情報に対して正または負の評価を与えるニューラルネットワークを持ち、その結合重みの初期値が遺伝子として与えられる。各エージェントは行動を出力するネットワークに従って行動するが、その際評価基準を与えるネットワークの評価を強化信号として、バックプロパゲーション法により行動を出力するネットワークの結合重みを更新するものとした。

その結果、進化の過程の初期段階においては、正しい評価を与えるネットワークを持つ個体がうまく学習ができることで多く生き残り、続いて、次第に学習を必要とせず先天的に正しい行動をとる個体群へと進化し、Baldwin 効果が確認されたと報告している。また彼らは、Baldwin 効果が有効に働いた後、評価基準を構成していた遺伝子は不要になって適応度に影響を与えなくなり、遺伝的浮動の力を大きく受けるようになったとも報告しているが、これは進化と学習の相互作用における相反的な側面のひとつであるとも考えられ、興味深い。この研究は、学習に関してより具体的なモデルを用いて、Baldwin 効果を示した点で重要であると言える。また、強化学習の分野において、学習の効率に大きな影響を与える強化信号をいかにして与えるかが問題となっているが、ERL は強化信号を与える規準自体もエージェントに内包させ、進化の枠組みに取り込んでいる点で興味深い。

2.1.3 動的環境における進化と学習の相互作用

ところで、これまで Baldwin 効果に関して行われた進化実験は、前述のように世代を通して最も適応的な遺伝子構成が不変であるという意味で静的な環境を設定したものであり、動的な環境において Baldwin 効果がどのように働くかについての議論は不十分であった。しかし、Baldwin 効果のような進化と学習の相互作用が起きるために、環境が静的であることは、必ずしも必要ではない。なぜなら、Baldwin 効果の働く要因である学習によるメリットとコストは、動的環境においても存在するからである。たとえば、刻々と変化する状況に対し、学習は柔軟に振る舞う術を与えてくれる。また、学習は常に良い結果ばかりをもたらすわけではなく、環境の変動が誤った学習や不必要な学習を引き起こす状況も考えられる。したがって、動的な環境においても、学習のメリットやコストが働いて Baldwin 効果のような現象が起きることは十分期待できる。むしろ、現実世界において、学習は不安定な環境に柔軟に適応するために不可欠であることを考えると、動的な環境での議論はより現実的あり、議論すべき重要な状況設定であると言える。

動的な環境を考える場合、大きく 2 つに分けることが出来る。一つは集団が置かれた環境自体が世代を通して変化し、集団中の個体の適応度に影響を与える場合である。Anderson は、世代を通して最適解が変動する動的な環境において学習が遺伝的な流れに与える影響を、力学系の方程式を用いて定量的に分析した。その結果、動的な環境においては学習にコストがかかっても学習に依存する集団へと進化することを示した[Anderson 95]。彼はまた、個体の学習には遺伝的な多様性を維持する働きがあり、これが環境の変動に追従することを可能にしているとも主張している。

Sasaki らは、ニューラルネットワークを用いて学習する個体が、世代を通して変化する食べ物または毒を表すビット列の入力を正しく識別するようにニューロンの結合重みを変更し、適応度に応じて進化するというシミュレーションを行った[Sasaki 99]。このとき、学習結果（学習後の結合重み）が次世代の個体に遺伝する Lamarck 型の進化システムと、学習結果が遺伝せず学習前の初期値が遺伝する Darwin 型の進化システムを用いて実験を行ったところ、Lamarck 型の進化システムでは動的環境に追従しきれなかったが、Darwin 型の進化システムでは進化の過程で学習を前提として次第に動的環境自体に適応していったと報告している。

また、Yamasaki らは、変数の入力に対する正しい出力をニューラルネットワークによって学習する個体群を、Darwin 型または Lamarck 型のシステムを用いて進化させ、学習すべき関数を動的に変化させた場合における双方のシステムの挙動について解析した[Yamasaki 00]。このとき、変動の幅が大きく離散的な環境においては、Lamarck 型よりも Darwin 型の進化システムの方がよい結果となることを示した。これらは、動的な要因が環境自体にある条件での進化と学習の相互作用に関する議論

の重要性を示したものである。

もう一つは動的な要因を集団の各エージェント自体が内包しているような場合である。このような状況として、例えば、集団における各個体の適応度が、個体間の相互作用によって決定されるような状況が考えられる。これまで、前者に注目した進化と学習の相互作用についての研究は前述のとおりいくつかなされてきたものの、後者についての Baldwin 効果に関する研究はほとんどなされていなかった。

2.1.4 目的

そこで、本研究は、個体間相互作用に依存した動的な環境において、進化と学習の相互作用である Baldwin 効果がどのように働くかについて知見を得ることを目的とする。このために、各個体の適応度が集団における個体間の相互作用に依存して決定され、世代を通して最適な解が決定できないという意味で動的な環境として、ゲーム論的状況設定を採用する。具体的には、利己的集団における協調行動の創発に関する抽象モデルである、繰り返し囚人のジレンマゲームにおける戦略の進化モデルを取り上げる。ゲームにおける決定論的戦略に対し、表現型可塑性に基づく学習の枠組みを導入することで、協調行動の進化に対する学習の影響について明らかにする [鈴木 99a, 99b, 99c, 99d, 00a, 00b], [Suzuki 00a, in press], [Arita 00]。したがって、今回焦点を当てる表現型可塑性は、個体の行動に関する可塑性 (behavioral plasticity) に属するものである。特に、本研究では、次のような点に注目して、実験および議論を行う。

- 戦略に学習を導入することで、個体間相互作用に基づく動的な環境においても Baldwin 効果は確認できるだろうか？
- 囚人のジレンマゲームに関する議論の焦点のひとつである、戦略集団における協調関係の創発に、学習はどのような影響を与えるか？
- その結果現れる戦略とはどのようなものか？
- これらの進化と学習の相互作用は、遺伝子の記述などの進化メカニズムや環境条件の違いによってどのような影響を受けるか？

なお、本研究で扱うような学習とは異なるが、個体間相互作用に基づく遺伝的でない要因によって個体の適応度が決定されるような系に関する研究もなされている。Watson らは、個体間の共生関係が進化の過程に与える影響について議論している [Watson 99a]。彼らは、前述の Hinton と Nowlan のモデルに対し、遺伝的に未決定な形質 “?” の値を決定する際、任意に取り出した他の個体における同じ遺伝子座の遺伝子を参照し、それが正しければその値を採用するという変更を行った。彼らは、Hinton らのモデルにおける学習プロセスを、各個体が生涯のうちに行う化学反応プロセスと見なし、“1” をそのプロセスを触媒する方向への正しい効果をもたらす遺伝子、“0” を正しくない効果をもたらす遺伝子、“?” (Watson らの論文では “-”)

をそのどちらでもない中立な遺伝子であるものとした。このとき、中立な遺伝子を他の個体の正しい値に置き換えることは、環境に共存する他の個体からも触媒を受けることに相当する。このため、従来のモデルでは、適応度がその個体の遺伝子構成のみに影響していたのに対し、変更したモデルでは、他の個体の遺伝子構成にも影響されることになる。

進化実験の結果、Baldwin 効果と同様に、次のような2段階の進化の過程が確認された。はじめに、中立な遺伝子を相互に触媒し補完し合うことで適応的となる遺伝子構成を持った個体群が集団中を占めた。これが、このモデルにおける共生関係の成立である。その後、共生関係を築くことができるグループの大きさに制限がある（大きさがその頻度に反比例する）場合には、グループの大きさが減少すると同時に“1”の遺伝子の割合が増加し、最終的には1個体で全て正しい遺伝子構成を持った個体が出現した。彼らは、このように、部分的に適応的な個体間において共生関係が築かれた上で、その適応的な部分が徐々に集積されていく過程を、*symbiotic scaffolding* と呼んでいる。

また、Kaneko らは、個体の発生過程における表現型の可塑的变化が同所的種分化を促進するという興味深い現象を、抽象モデルを用いて示している[Kaneko 00]。彼らは、まず、発生プロセスの初期状態と個体間相互作用の決定する遺伝子構成が全く均一の集団においても、発生過程における化学物質のやり取りに関する相互作用が集団内の表現型の分化を引き起こすことを示した。そのような状況を前提にして集団の遺伝子構成が進化した場合、分化した表現型に対し1対1対応を持つように遺伝子構成も徐々に分化し、種分化が確認されたと報告している。これらの研究は、本研究で扱う学習とは異なるものの、個体間相互作用による適応度の変化が進化の過程にもたらす興味深い現象を示している点で重要だと言える。本研究においても、学習する個体同士がもたらす相互作用が、進化の過程に興味深い影響を与えることが期待される。

2.2 表現型可塑性を導入した繰り返し囚人のジレンマゲームにおける戦略の進化モデル

以上を踏まえ、個体間の相互作用のみを考慮した動的環境として、遺伝的アルゴリズムを用いた繰り返し囚人のジレンマゲームの戦略の進化モデルを構築した。繰り返し囚人のジレンマゲームについて説明した後、モデルについて解説する。

2.2.1 繰り返し囚人のジレンマゲーム

繰り返し囚人のジレンマゲームは、2人非ゼロ和ゲームの一種であり、Axelrodによる戦略のコンピュータトーナメントの研究[Axelrod 84]をはじめとして利己的集団における協調行動の創発に関して数多くの研究がなされている。ゲームは表 2.1 に代表される利得行列を用いて以下の手順で行われる。

- 2人のプレイヤーは協調 (Cooperate) または裏切り (Defect) のどちらかの手を同時に出す。
- 出した手に応じて、利得行列 (表 2.1) から両者が得る得点が決まる。
- この対戦を繰り返し行い、その合計 (平均) 得点を競う。

表 2.1: 囚人のジレンマゲームの利得行列

相手の手 (→) 自分の手 (↓)	協調 (C)	裏切り (D)
協調 (C)	($R=3, R=3$)	($S=0, T=5$)
裏切り (D)	($T=5, S=0$)	($P=1, P=1$)

(自分の得点, 相手の得点)

$$2R > T+S$$

1回きりの対戦において、プレイヤーがともに自らの期待利得を最大にするような戦略、つまり支配戦略である裏切り (D) を取った場合、裏切り合いとなりこれはナッシュ均衡解である。にもかかわらず、この解はパレート最適ではなく、双方にとってより良い解すなわち協調し合いが存在するため、裏切りは正しい判断ではなかったのではないかというジレンマが生じる。さらに、十分長い繰り返しゲームにおいては、交互に裏切るよりも協調し合ったほうが双方の利益となる ($2R > T+S$) ため、いかにして協調関係を築くことができるかが高い得点を得る際の問題となるが、協調関係を築くことができるかどうかは相手の出方次第である。つまり、ゲームにおいてある戦略がうまくやれるかどうかは、対戦相手に大きく依存するのである。したがって、ジレンマゲームの戦略集団における総当たり戦の得点を適応度とするような環境を考えると、各個体の適応度が他の個体の振る舞いに依存し、世代ごとに

刻々と変化する動的な環境として捉えることができる。

2.2.2 戦略の遺伝子表現

集団における各個体は繰り返し囚人のジレンマゲームの戦略を遺伝子として持つ。このモデルでは、各個体の持つ戦略を戦略遺伝子列 GS と可塑性遺伝子列 GP の2つの遺伝子列の組で表現する。戦略遺伝子列は Lindgren のモデル[Lindgren 91]と同様な、過去の対戦履歴に依存して次回の手を決定する決定論的戦略を定義する。記憶長 m の戦略は裏切りを“0”，協調を“1”として以下のような2進数で表された対戦履歴 h_m を持つ。

$$h_m = (a_{m-1}, \dots, a_1, a_0)_2 \quad (2.2)$$

ここで a_0 は前回の相手の手, a_1 は前回の自分の手, a_2 は前々回の相手の手…とする。

ある履歴 x に対応して次回出すべき手を A_x (“0”：裏切り, または“1”：協調) とすると, 記憶長 m の戦略は,

$$GS = [A_0 A_1 \dots A_{n-1}] \quad (n = 2^m) \quad (2.3)$$

と表すことができる。これを戦略遺伝子列とする。さらに, 各 A_x に対してその表現型が可塑性を持つかどうかを P_x (“0”：可塑性を持たない, “1”：可塑性を持つ) として, 可塑性遺伝子列を,

$$GP = [P_0 P_1 \dots P_{n-1}] \quad (2.4)$$

と定義する。例えば, しっぺ返し戦略(初回は協調, 以降は前回相手が出した手を真似る) [Axelrod 84]を記憶長2で表すと, $GS=[0101]$, $GP=[0000]$ となる。

2.2.3 学習規則

可塑性を持つ表現型を対戦中に用いた場合, その表現型は学習規則によって変更される。本研究では2つの学習規則を取り上げる。一つはランダム型学習, もう一つは学習行列を用いたメタ・パブプロフ学習である。学習は次の手順で行われる。

- 繰り返し対戦を行う前は, 各個体は GS の表す決定論的戦略をそのまま表現型として持つ。
- 表現型と履歴を参照し対戦を行い, 用いた表現型 (C または D) に対応する

可塑性遺伝子列のビットが“1”（可塑的）であった場合、以下のどちらかの手順で用いた表現型を書き換える。

1. ランダム型：その表現型をランダムにCまたはDと置き換える。
 2. メタ・パブプロフ学習：その表現型を、対戦結果に対応するメタ・パブプロフ学習行列の値（CまたはD）と置き換えたものを新たな表現型とする。
- 次回の対戦以降、新たな表現型を参照し、手を決定する。

表 2.2:メタ・パブプロフ学習行列

相手の手 (→)	C	D
自分の手 (↓)		
C	C	D
D	D	C

ランダム型学習は、対戦の結果によらずに表現型を変更する点で学習とは言い難いが、本研究では表現型の可塑性を学習の根本として捉えていることから、最もプリミティブな学習規則の一つとして取り上げる。

メタ・パブプロフ学習では学習行列の値として表 2.2 に示す行列を定義し、これに基づいて表現型を変更する。この行列の値は、ゲームを行った結果得られる得点が相対的に高ければそのまま変更せず、逆に小さければ変更するという強化学習の原理に基づくものであり、学習規則としてシンプルかつ典型的なものとして今回採用する。この行列自体はパブプロフ戦略（初回は協調、以降は対戦結果が相対的に良ければ次回も同じ手を出し、悪ければ手を変える）[Nowak 93]と同じであるが、直前の対戦結果に応じて次回出す手を決定するのではなく、表現型を用いた結果に応じて戦略自体を変更（学習）するという意味で、この行列を用いた学習方式をメタ・パブプロフ学習と呼ぶ。

ここで、メタ・パブプロフ学習の例として、 $GS=[0001]$ 、 $GP=[0011]$ の戦略が学習する例を示す。図 2.3（学習前）はこの戦略の表現型を図示したものである。記憶長 2 の履歴（前回の自分の手と相手の手）に対応して、次回出す手が表現型として決められている。ただし可塑性を持つ表現型には下線を引いた上で、初期状態を示している。過去の対戦履歴が CC であったと仮定すると、表現型からこの戦略は C を出す。このとき相手が D を出したと仮定する。ここで、C を出すのに用いた表現型は可塑性を持つのでメタ・パブプロフ学習行列をもとに表現型を変更する。この場合、対戦結果は自分の手が C、相手の手が D なので、学習行列から表現型を D に変更し、次回以降の対戦履歴が CC の場合には D を出すようになる。従って、戦略の表現型は図 2.3（学習後）のように変化するということである。このように、可塑性遺伝子列に 1 のビットを持つ戦略個体は、繰り返し対戦を通して表現型が変化しうるという意味で可塑的な戦略であると捉える。

対応する学習遺伝子のビットが 1 である戦略遺伝子の値は表現型の初期値としてのみ働く。そこで、今後各戦略を、可塑性を持つ学習遺伝子に対応する戦略遺伝子を x と置き換えた戦略遺伝子列でまとめて表現することで、進化の過程の大枠を捉えることにする (e.g. $GS=[1000]$, $GP=[1001]$ ならば $[x00x]$ と表す)。なお、本研究では、学習すること自体にかかる明示的なコストを導入せず、学習にかかるコスト (及びメリット) はすべて個体間の相互作用の結果として与えられるものとする。

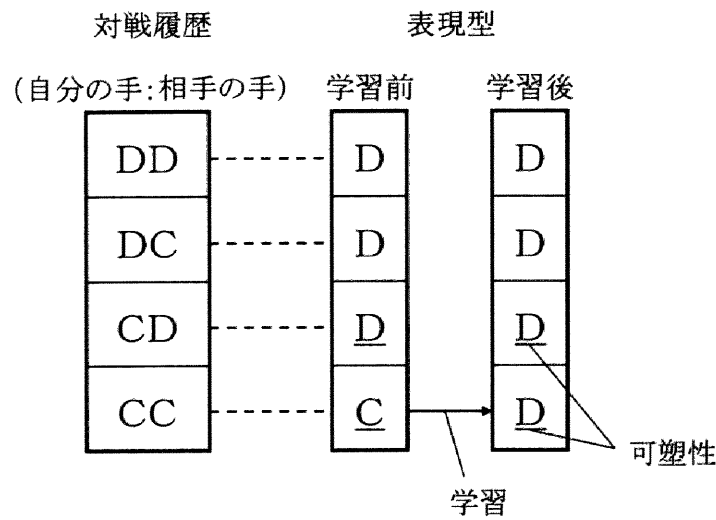


図 2.3:メタ・パブロフ学習の例 ($GS=[0001]$, $GP=[0011]$)

2.2.4 繰り返し対戦と進化

以上のような戦略個体同士でノイズありの繰り返し対戦を表 2.1 の利得行列を用いて行う。ノイズとは、繰り返し対戦において、各個体が表現型に従って出すべき手が確率 p_n で反転してしまうことで、現実世界における表現の間違い、転送経路のノイズ、誤解などの不可抗力を象徴するものである。また、2.2.2 節で示したとおり、本研究で用いられる戦略が手を決定するためには履歴が必要である。そこで、各繰り返し対戦の初回は、繰り返し対戦ごとにランダムに作成された仮想の履歴を各個体が参照して各自の手を決定するものとする。本モデルでは、繰り返しの回数は固定せず、対戦ごとに確率 p_d で次回の対戦が行われるものとする。これは、十分長い間繰り返されるが、実際何回繰り返して行われるかはプレイヤーには分からないという設定を反映したものである。したがって、繰り返し対戦回数が n である確率は $P(n)=(1-p_d)p_d^{n-1}$ ($n=1, 2, \dots$)となる。また、可塑的な戦略における表現型は、繰り返し対戦ごとに初期状態 (戦略遺伝子列が示すままの状態) に戻されるものとする。

このような繰り返し対戦を集団 (個体数 N) において総当たりで行い、その合計得点を各戦略個体の適応度とする。最後に、各適応度に比例したルーレット選択により次世代の集団を生成する。その際、各遺伝子について確率 p_m でビットが反転す

る突然変異を導入する。なお、計算量を軽減するために、繰り返し対戦において、2 個体の遺伝子情報の組み合わせが初めて行われるものである場合は、繰り返し対戦を 20 回行ったときの平均得点を対戦結果として用いるとともにその得点を保存し、既に同じ組み合わせによる対戦が行われていた場合には対戦を行わず、保存していた平均得点を対戦結果として用いる。ただし、500, 1000, 1500 世代ごとに、それまでに保存していたすべての得点を消去し、新たに計算し直すものとする。

2.3 基本的な進化実験

以上のモデルを用いて行った進化実験の結果について示す。はじめに、基本的な実験として記憶長 2 の戦略群を用いて、学習なしの場合、学習を導入した場合について実験を行った。

2.3.1 学習なしの戦略での実験

戦略に学習を導入しない場合の集団の挙動を観察するために、各戦略の可塑性遺伝子列 GP を常に[0000]に固定した上で、戦略遺伝子列 GS の各ビットをランダムに選んだ記憶長 2 の戦略群を初期状態として、進化実験を行った。なお、パラメータとして $p_m=1/1500$, $N=1000$, $p_r=1/25$, $p_d=99/100$, 世代数 2000 を用いた。今後、特に断りのない限り、各実験においてこれらのパラメータを用いる。

進化実験の典型的な結果を図 2.4 に示す。世代を通して塗りつぶされた領域は集団中を占めた遺伝子の分布を示す。また平均得点（白実線）は各世代に行われたすべての対戦の得点を平均したもので、互いに協調し合ったときに最も高くなる（3 点）ことから協調の度合いを表す指標として捉える。同図より、[0101]（しっぺ返し戦略）→[1001]（パブプロフ戦略）→[0001]→[0101]→…というサイクルに代表されるような、裏切りのない戦略と協調的な戦略が交互に集団中を占める状態が続き、平均得点は大きく振動することがわかる。この結果は Lindgren の報告した進化の過程と類似したものである。

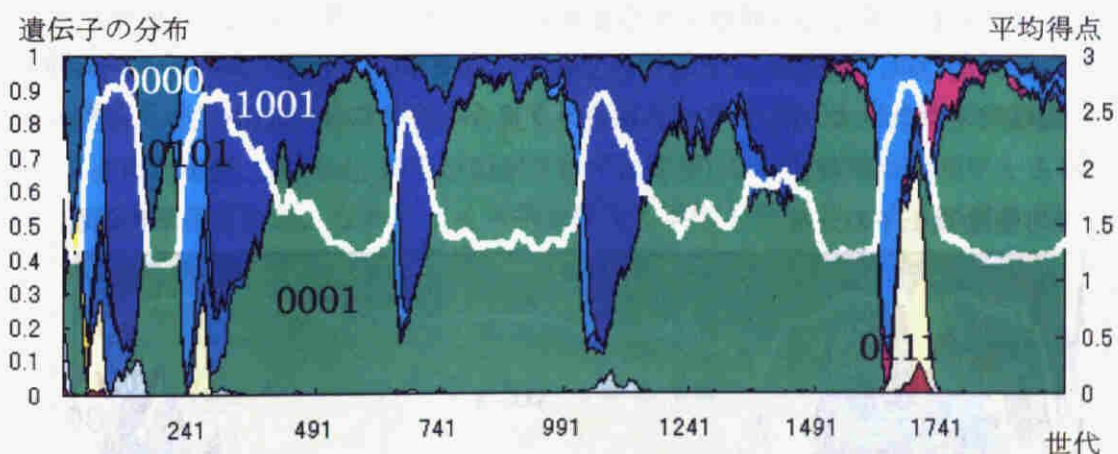


図 2.4: 記憶長 2 学習なしでの実験結果

2.3.2 記憶長 2 ランダム型学習

次に、記憶長 2 の集団（初期集団における GS 、 GP 各遺伝子の値はランダムに決定するものとする）において学習方式をランダム型学習として同様のパラメータを用いて進化実験を行った。実験結果の典型的な例を図 2.5 に示す。ただし、試行ごとに進化の挙動は大きく異なることを断っておく。ここで集団の可塑性（黒実線）とは、可塑性遺伝子列中に占める 1 のビットの割合を示し、これは 2.1 節における学習依存度に相当する指標として捉えることができる。

この試行では、初期集団から裏切りの戦略が集団中を占めて平均得点が低下した。つづいて若干の集団の可塑性の増加を伴いながら $[x0x1]$ や $[01x1]$ という戦略が集団中の大半を占めて平均得点の高い集団へと進化した。これらの戦略は協調し合いを維持するが、ノイズなどにより協調関係が崩れると、その可塑性によって確率的に協調関係を回復する戦略である。その後、可塑性は低下し、 $[x001]$ 型の戦略が長い世代にわたって集団中を占める状態が続いた。この戦略は、同種同士の対戦において基本的に協調し合い、ノイズが入って協調関係が崩れると、可塑性により確率的にはあるがより早い段階で協調し合いに戻ることを出来る戦略である。しかし、集団はこのまま収束する訳ではなく、さらに可塑性は低下し、より協調的な $[1001]$ 型の戦略が徐々に侵入してきた。この $[1001]$ 型戦略が増加する方向への進化は集団全体の性格をより協調的に変化させるため、裏切りの $[0001]$ 型戦略の侵入を許し、平均得点の低下を招いた。

以上のうちで、集団が $[x001] \rightarrow [1001] \rightarrow [0001]$ のように進化する過程は他の試行においても多く確認された。 $[x001]$ 型戦略は他の試行においても集団中を長い期間占めたことから、比較的安定的な戦略であると言える。しかし、集団が完全に収束するに至らなかったのは、記憶長 2 ランダム型学習の集団では、学習メカニズムが貧弱で、可塑性のメリットとコストのバランスがうまく取りきれないからであると考えられる。

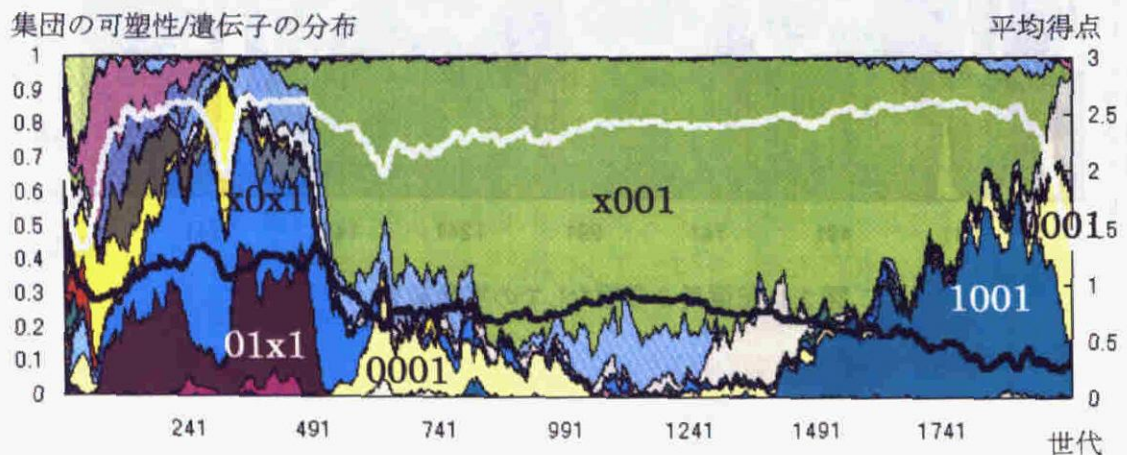


図 2.5: 記憶長 2 ランダム学習での実験結果

2.3.3 記憶長 2 メタ・パブプロフ学習

次に、記憶長 2 メタ・パブプロフ学習の集団（初期集団における *GS*, *GP* 各遺伝子の値はランダムに決定するものとする）において同様のパラメータを用いて進化実験を行った。

実験結果の一例として図 2.6 (2000 世代), 2.7 (300 世代) を示す。この試行における進化の過程の概略を述べる。はじめの約 60 世代までは、裏切りの戦略 ([0000], [000x] など) が平均得点を低下させた。またそれとほぼ同時に [0x0x], [00xx], [0xxx] などの一部可塑的な戦略も集団中を占めた。その後、集団の可塑性の増加とともに [xxxx], [x0xx] といった可塑的な戦略が集団中を占めて協調関係を築き、約 250 世代までに高い平均得点を持つ集団へと進化した。これまでの過程で、可塑性は裏切りの集団から協調的な集団へのシフトに有利な方向へと働いたと考えられ、これは Baldwin 効果の第 1 段階と捉えられる。

その後、高い平均得点を維持したまま、集団の可塑性は次第に低下して約 50% のところで安定し、最終的には集団の大部分を [x00x] 型の個体が占める結果となった。これは、一部の可塑性がコストとして働いて、集団を維持するのに最低限必要な可塑性を持った戦略が選択されたためと考えられ、Baldwin 効果の第 2 段階と捉えられる。

行った試行のほとんどで、最終的に [x00x] 型の戦略が集団中を占めた。また、この結果に見られるような傾向をもつ進化の過程が、試行の約 70% で確認された。試行の残り 30% の中には、第 1 段階の後、可塑的な集団から第 1 段階の初期の裏切りの戦略が集団中を占める状態に戻ることが何度か繰り返される場合があった。これは、過度に可塑的であることが裏切りの戦略の侵入を許した結果、[x00x] への進化を遅らせたことがこの原因と考えると、進化と学習の相反的側面と捉えることができ興味深い。このほか、初期段階から直ちに [xxxx] などの可塑的で協調的な戦略が集団中を占めた場合、Baldwin 効果が観察されずに直接 [x00x] 型戦略が集団中を占めた場合などが確認された。なお、ノイズ率を変えて実験を行うと、ノイズ率が高いほど裏切りの戦略（特に [0001] 型戦略）が集団中に広まりやすい傾向が見られた。

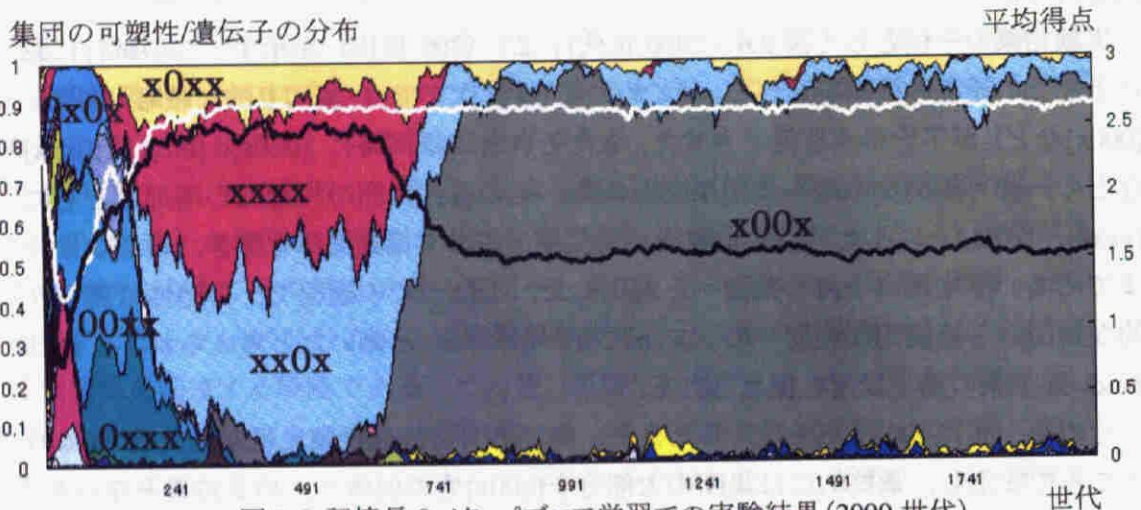


図 2.6: 記憶長 2 メタ・パブロフ学習での実験結果 (2000 世代)

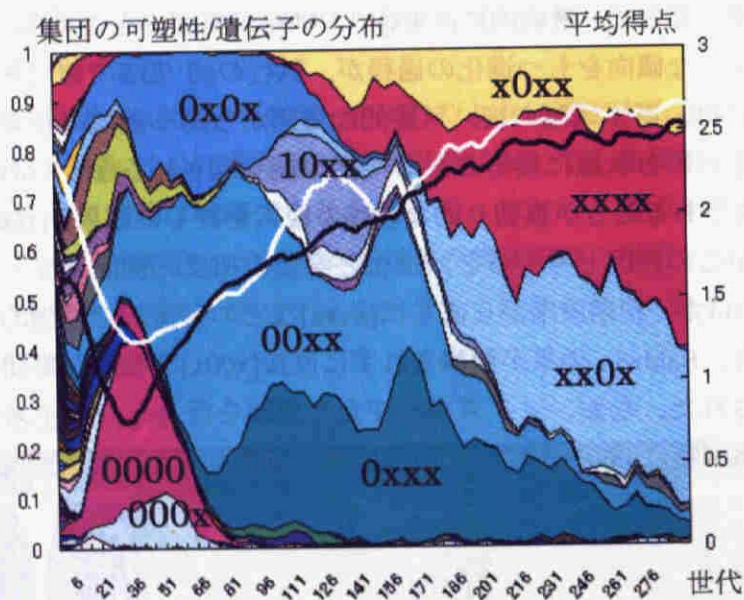


図 2.7: 記憶長 2 メタ・パブロフ学習での実験結果 (300 世代)

2.3.4 戦略の推移と Baldwin 効果

これまでの結果から、このような動的な環境においても Baldwin 効果と捉えられる進化の過程が確認されたが、学習は戦略集団の進化に具体的にどのような影響を与えたのだろうか。前節の記憶長 2 メタ・パブプロフ学習の実験における戦略の推移と Baldwin 効果の関係について詳細に解析する。

集団の可塑性と平均得点の相関に注目して進化の過程を観察すると、Baldwin 効果の 2 つの段階をはっきりと把握することができる。図 2.8 は 10 回の試行における、集団の可塑性と平均得点の相関の軌跡を表したものである。初期集団の状態から、一旦平均得点と集団の可塑性が低い状態へと進化した後、共に上昇する方向（グラフ右上）へと進化しているのがわかる。これが Baldwin 効果の第 1 段階である。その後、軌跡はまっすぐ左へと向きを変え、平均得点を維持したまま、集団の可塑性のみが低下しているのがわかる。これが第 2 段階である。

この相関図の軌跡に合わせて、出現した戦略を大まかに分類すると、図 2.9 のようになる。はじめに、初期集団においては裏切りの戦略に対して搾取される戦略が多く含まれるため、相関図左下のような全面裏切りの戦略が有利となり、集団中に広まる。左下の裏切りの戦略群を中心とした戦略が集団中を占めると、対戦が裏切り合いばかりになり平均得点は低下する。この状態では、右下のような戦略が徐々に集団中に広まる。これらの戦略は左下のような戦略と同様に一度裏切り合っ

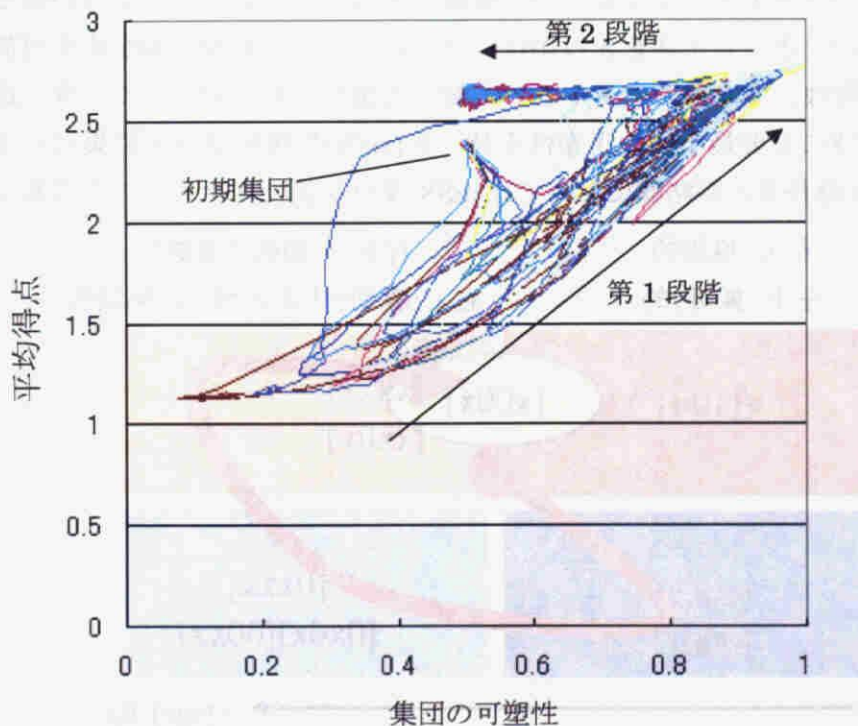


図 2.8: 進化の過程における集団の可塑性と平均得点の相関

たら裏切り合いを続けるが、同種同士の対戦では戦略に含まれる可塑性によってノイズや初期状態をきっかけにして協調し合うので、裏切り合いと比べて若干高い得点を得るためである。

相関図右下のような戦略が増加し、集団全体においてさらに協調する機会が増え、平均得点が上昇してくると、裏切りの戦略に対してそれほど点を与えないことを維持しつつも、右下の戦略と比べてより協調的であることで高い得点を得る、右上のようなより可塑的な戦略が集団中を占めるようになる。ここで、最終的に集団の大半を占める[x00x]型戦略よりも右上の戦略群の方へと進化する傾向が高いのは興味深い。左下から右下、右上の戦略群への進化のように平均得点が上昇している途中では、他種の得点を下げるよりも、まず、自身の得る得点が高いことが集団中に広まるために必要とされる。しかし、[x00x]のESS的な性質(2.4.1節で後述)が右下の戦略群との対戦で両者の得る得点を右上の戦略群同士の場合と比べて下げてしまっているため、右上の戦略が先に集団中を占めると考えられる。このことは、ESS的な戦略であるからといって、他の戦略集団に容易に入り込むことができる訳ではないことを示している。これまでの可塑的な協調集団に至る過程が Baldwin 効果の第1段階である。

その後、右上のような協調的な戦略ばかりになると、対戦のほとんどが協調し合いになり、異なる戦略同士の適応度の差が小さくなる。この状態においてはほとんどの対戦は基本的に協調し合い、ノイズが入ると一定の回復過程を経て再び協調し合いに戻るというサイクルになるため、戦略の差はノイズが入ってから振る舞いに現れる。このとき、ノイズをきっかけに自分にとって不利な協調を出す可能性のある余分な可塑性、すなわちコストとして働く可塑性が取り除かれていき、最終的には協調の維持に必要な最小限の可塑性を持った[x00x]型戦略が徐々に集団中を占める。[x00x]は記憶長2の戦略群においてはESS条件を満たすため、ここで集団は安

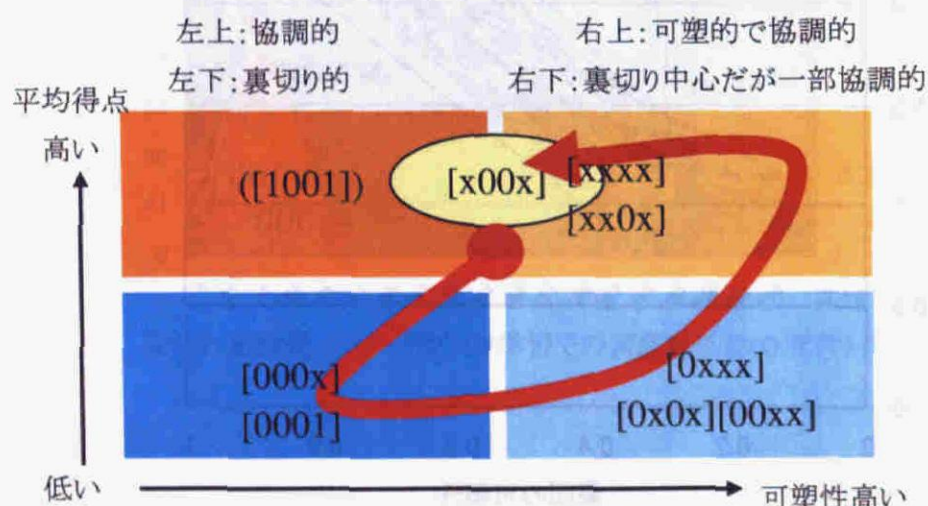


図 2.9: 集団の可塑性・平均得点と進化の過程で出現した戦略の分布

定する。この過剰な可塑性の減少と安定化が、Baldwin 効果の第 2 段階である。

以上から、この過程において、学習は裏切りの戦略集団から協調的な戦略集団へのシフトと安定化に大きな影響を与えていると考えられる。また、戦略における可塑性は、対戦する戦略に応じて自身の振る舞いをうまく変える特徴を実現するのに有効であると考えられる。

2.4 メタ・パブプロフ[x00x]型戦略の解析

これまでの結果から、最終的に集団中を占めたメタ・パブプロフ[x00x]型戦略は非常に安定していることがわかった。そこで、この戦略の安定性や特徴、学習の果たす役割について、いくつかの点から解析する。

2.4.1 ESS 条件

集団における戦略が進化の過程で安定であるかどうかの基準として、Maynard-Smith が提案した「進化的に安定な戦略 (ESS)」[Maynard-Smith 82]がある。集団において ESS を満たす戦略 a の条件は、

$E(a, b)$ を戦略 a と b との対戦で a が得る得点として、

$$E(a, a) > E(b, a) \quad (2.5)$$

または

$$E(a, a) = E(b, a) \quad \text{かつ} \quad E(a, b) > E(b, b) \quad (2.6)$$

が他のすべての種類の戦略 b に対して成り立つことである。

メタ・パブプロフ[x00x]型戦略がこの条件を満たすかどうか確認するために、[x00x] ($GS=[0000]$, $GP=[1001]$) と記憶長 2 の可能なすべての戦略 256 個との繰り返し対戦を、 $p_n=1/25$, $p_d=99/100$ で 100 回行ったときの各対戦の平均得点を計算した。図 2.10 は[x00x]と記憶長 2 の戦略との対戦成績である。横軸は、各戦略の遺伝子列を[GSGP]とならべて 8 ビットの 2 進数として見た場合の値を表す。縦軸は、[x00x]同士の対戦で[x00x]が得た得点と各戦略個体と[x00x]との対戦において各戦略個体の得た得点の差、すなわち (2.5) 式の左辺と右辺の差である。したがって、[x00x]が ESS であるには、同種同士の対戦を除いて縦軸の値がすべて 0 より大であれば良く、この図から ESS 条件を満たしていると言える。

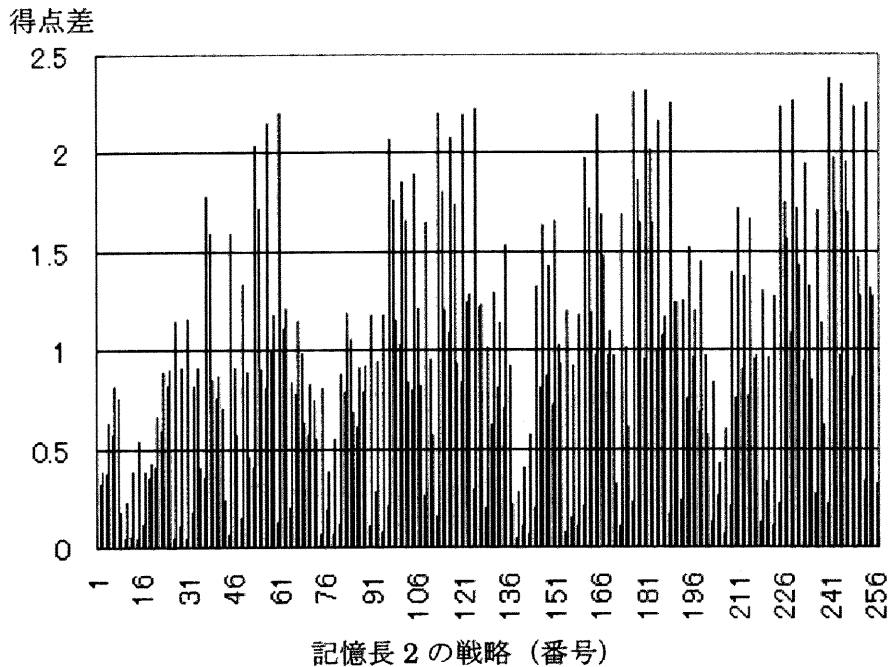


図 2.10: メタ・パブロフ[x00x]型戦略と他の記憶長 2 の戦略との対戦成績

2.4.2 状態遷移図に基づく解析

メタ・パブロフ[x00x]型戦略の振る舞いは、2つの可塑性を持った表現型と、現在の対戦結果を内部状態とすることで、状態遷移図として表現できる。図 2.11 に[x00x]型戦略の状態遷移の様子を示す。各矩形は自分と相手の出した手、および可塑性を持った表現型の状態を示す。上下に並んだ“0”または“1”は、上が相手の出した手、下が[x00x]の出した手を表す(“0”:裏切り, “1”:協調)。[x00x]の出した手につけられた添え字は、可塑性を持った表現型の現在の状態を示し、順に[x00x]の前のxと後ろのxの表現型の状態を示す。各状態から2本ずつ伸びた矢印は、相手の取りうる手に依存して可能な状態遷移先を示している。任意の戦略と[x00x]との繰り返し対戦は、相手の戦略に応じて矢印を選んでいくことで表現できる。たとえば、[x00x]と全面裏切り戦略[0000]が対戦した場合、相手の手が常に0となるように矢印を選んでいけばよい。この場合、たとえばサイクル②を取り続けることになる。

この図において注目すべき点は、協調関係を持続する状態(状態 A)が崩れたとき、もう一度もとの状態 A に最短で戻るには、相手が「裏切り, 協調, 裏切り」という複雑な手(サイクル①)を取らなければならないことである。実は、「裏切り, 協調, 裏切り」のあと協調に戻るという協調の回復過程は、[x00x]同士の対戦において実現される。つまり、協調関係が崩れたときに最も早く協調関係を回復することのできる相手の一つが同種であるということである。逆に言えば、同種以外の戦略

に対して、 $[x00x]$ は協調関係を回復しにくいことを表しており、この特徴は $[x00x]$ のESS的な性質を生み出すメカニズムの一つであると見なすことができる。

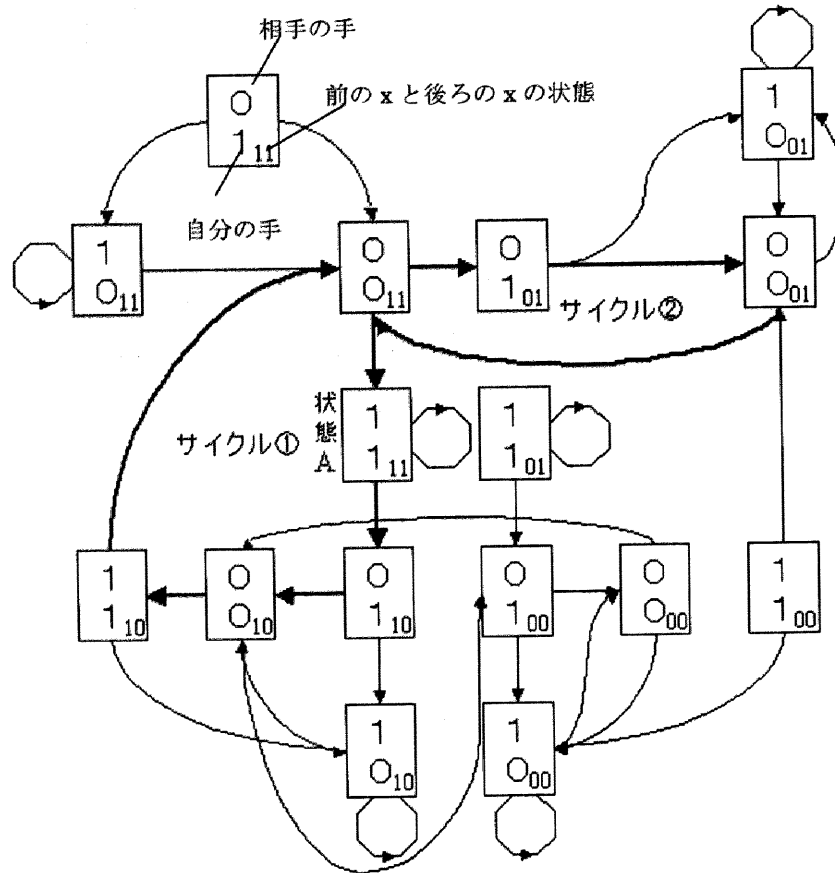


図 2.11 メタ・パブプロフ $[x00x]$ 型戦略の状態遷移図

2.4.3 学習の役割

これまで、進化の過程において、過剰な可塑性が減少した結果、最終的に[x00x]型の戦略が集団中を占めたと論じてきたが、ではどうして、「前の x」と「後ろの x」という 2 つの可塑性を残して減少は止まったのだろうか。状態遷移図からわかるとおり、戦略の特徴は遺伝子列全体を総合して決まるが、ここではあえて前の x、後ろの x、その両方それぞれの役割について検討する。

前の x は裏切り中心の戦略に対して、多く点を取らせず、侵入させないという点で有効であると考えられる。図 2.11 において、サイクル②は全面裏切り戦略と対戦した場合にたどるものの一つである。このとき、前の x の可塑性により[x00x]は 2 回の裏切り合いの後 1 回裏切られる。[x00x]の得る得点は約 0.8 点と低いものの、このおかげで全面裏切り戦略が得る得点は約 2.3 点に下がる（表 2.3）。[x00x]同士の対戦で[x00x]が得る得点は約 2.6 点（ノイズの影響を含む）であるから（表 2.5）、これは[x00x]を進化的に安定な戦略にするのに貢献している。

表 2.3：全面裏切り戦略と[x00x]との対戦例

0000	000000000000…	平均約 2.3
x00x	010010010010…	平均約 0.8

一方、後ろの x は協調中心の戦略に対して、ノイズによる偶発的な裏切りをきっかけにして裏切りに転じる点で有効であると考えられる。図 2.11 中央の互いに協調し合った状態から、協調が崩れるとき、後ろの x は 1（協調）から 0（裏切り）に変わる。これは次回協調し合った後は裏切ることを示しており、これは協調を維持するような戦略に対して搾取する（表 2.4）とともに、相手に点を取らせない点で有効である。

表 2.4：[x001]と[x00x]の対戦例（下線はノイズ）

x001	1111 <u>00</u> 1101101…	平均約 1.6
x00x	1111101001001…	平均約 2.8

以上のとおり、前後の x にはそれぞれ裏切りの・協調的戦略に点を与えない働きがあると考えられる。しかし、この働きを生かすには、[x00x]同士の対戦で必ず強固な協調関係が築かれなければならない。前述の通り、[x00x]はサイクル①によってノイズに対しうまく協調を回復するが、この遷移の中では前後の x についてそれぞれ 2 回ずつ交互に学習（結果的に状態が変化しない学習も含む）が行われている（表 2.5）。これは前後の x による学習が共同してうまく働いた結果、強力な協調関係を築いていることを示している。

表 2.5 : [x00x]同士の対戦での協調の回復例 (下線はノイズ)

x00x	1111 <u>0</u> 0101111... 平均約 2.6
x00x	111110101111... 平均約 2.6

2.5 遺伝的記述の拡張

2.4 節では、基本的な記憶長 2 のモデルにおいて進化実験を行った。本章では、これまでのモデルに対し、記憶長の突然変異と学習行列の進化という 2 つの点から遺伝的記述の拡張を行うことにより、それぞれが創発的ダイナミクスにどのような影響を与えるかについて議論する。

2.5.1 記憶長の突然変異を導入した進化

これまでの実験では、各戦略は記憶長が固定されており、遺伝子列であらわすことのできる戦略の種類は限られていた。そこで、記憶長が長くまたは短くなる突然変異を導入することで、理論上戦略が制限されないオープンエンドな環境を設定し、進化実験を行った。

具体的には、2.3.3 節のメタ・パブプロフ学習を用いたモデルにおいて、遺伝的操作により次世代の個体を生成する際、各個体につき確率 1/3000 で *GS*, *GP* 各遺伝子列の長さを 2 倍（各遺伝子列について、遺伝子列を 2 つならべたものを新たな遺伝子とする）または 1/2 倍（各遺伝子列について、遺伝子列を 2 等分し、そのうちどちらかをランダムに選択して新たな遺伝子列とする）するものとした（例：2 倍[1101]→[11011101]、1/2 倍[1101]→[11]または[01]）。なお、この方法は Lindgren らのモデル [Lindgren 91] に基づいているが、記憶長が増加する方への突然変異において、彼らのモデルの場合は戦略自体の挙動は変化しないが、このモデルにおいては可塑性遺伝子列の影響により戦略自体の挙動も変化する点で異なる。

初期集団を、*GS* はランダム、*GP*=[00]とした記憶長 1 のメタ・パブプロフ学習を行う集団として行った典型的な結果を図 2.12 に示す。試行によってばらつきはあるが、

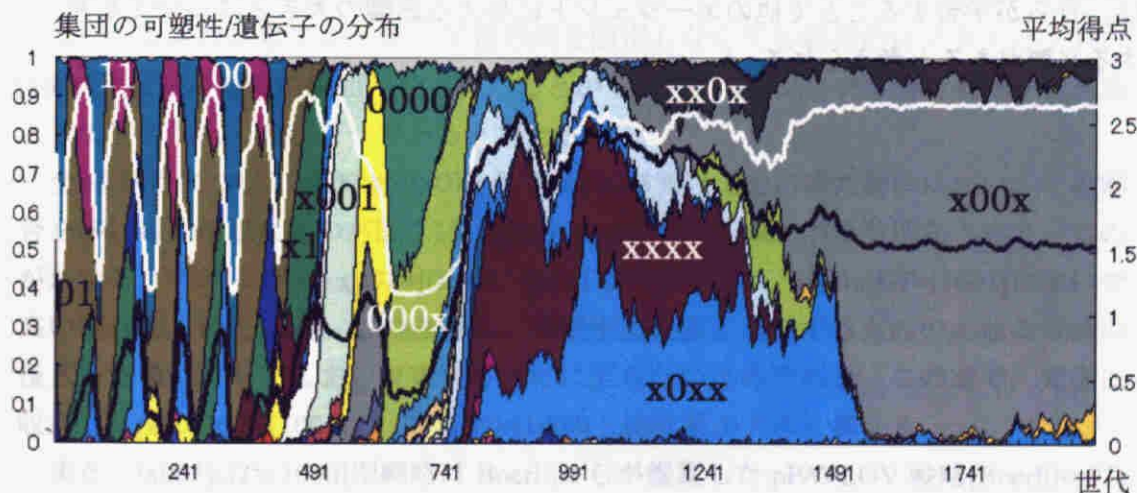


図 2.12: 記憶長の突然変異を導入した場合の実験結果

数百世代にわたって[01]（しっぺ返し戦略）など記憶長1の戦略が交互に集団中を占めたのち、記憶長1の集団から記憶長2の集団へと進化した。その後、試行の50%で2.3.3節での過程と同様な進化が見られ、それ以外ではBaldwin効果（集団の可塑性の増加と減少）は見られないまま、直接[x00x]型の集団へと進化した。前者の場合は可塑的で協調的な戦略（[xxxx]など）が現れる前に、裏切りの集団が集団中を大きく占め、平均得点を低下させている場合が多く、後者の場合、[x00x]型の戦略が現れる前に、すでに協調的な集団になっている場合が多く見られた。これは、集団の可塑性の上昇が、裏切りの集団から協調的な集団へのシフトとその後の[x00x]型戦略へのスムーズな進化に貢献しているのを支持するものと考えられる。

また、ごくわずかな場合において、メタ・パブプロフ[x00x]型戦略に収束せず、記憶長3以上の戦略が多種類集団中を占めることが確認されたが、一旦[x00x]型戦略の集団に進化した後は、それ以上記憶長の長い集団へと進化することはほとんど見られなかった。これは、メタ・パブプロフ学習を用いる場合には、記憶長2の戦略[x00x]が十分安定しており、それ以上記憶長の長い戦略は集団の安定のために必要とされなかったことを示している。

Tuomasらは、強化学習を行うマルチエージェント系において、各エージェント自体の学習が他のエージェントにとって不安定な環境を作りだしてしまうために学習がうまくいかず、結果として集団全体の利益を下げてしまう場合があることを指摘している[Tuomas 96]。メタ・パブプロフ学習は、強化学習の一つとして考えられるが、この結果は記憶長が長くなり学習が複雑になるほどよいというわけではないことを示している。実際、GS=[0000]GP=[1001]で表される[x00x]型戦略が記憶長の突然変異を起こしてできた個体（GS=[00000000]GP=[10011001]）同士が対戦を行うと、一方の学習による挙動の変化が他方の学習を結果的に邪魔して協調し合えない状態が続くことが観察された。したがって、マルチエージェント系における強化学習においては、自らが学習することで他のエージェントに与える影響の大きさについても考慮する必要があると考えられる。

2.5.2 学習行列を遺伝子として取り込んだ進化

これまでの学習行列を用いた実験では、行列の値としてメタ・パブロフ学習行列を定義し、前もって与えてきた。しかし、これとは異なる学習行列を持つ戦略がどのように振る舞うかは興味深い点である。そこで学習行列を表す遺伝子列 GT を新たに3つ目の遺伝子列として定義する。各個体はそれぞれ GT を持ち、 GT の表す学習行列を参照して学習を行うものとする。 GT は具体的に表 2.6 のように定義される(例：メタ・パブロフ学習行列: $GT=[1001]$)。

表 2.6：学習行列を表す遺伝子列 GT の定義

相手の手 (→)	C	D
自分の手 (↓)		
C	(CC)	(CD)
D	(DC)	(DD)

$$GT=[(DD) (DC) (CD) (CC)]$$

以上のような変更をモデルに加え、進化実験を 4000 世代に渡って行った。初期集団は遺伝子をランダムに決定した 100 個体をそれぞれ 10 個体ずつ用意するものとした。進化実験の結果を図 2.13, 2.14 に示す。図中の####:****はこれまでの遺伝子表記における記述が####, GT が****の個体であることを示す。なお、学習行列の進化を導入することにより、集団の可塑性は集団の学習に対する依存度を反映しにくくなるため、今回は考察には含めない。

60 回試行を行ったところ、そのうち 29 回で $[x00x]GT=[1001]$ すなわちメタ・パブロフ学習行列を用いた戦略が集団中を占めた (図 2.13)。また、そのうち 3 回で $[x001]GT=[1000]$ が集団中を占めた (図 2.14)。残りの 28 回はひとつの戦略には収束しない不安定な状態であった。学習行列を限定しなくても最終的にメタ・パブロフ $[x00x]$ 型戦略が頻繁に集団中を占めたことから、メタ・パブロフ学習行列は学習規則として妥当なもののひとつであると言える。

全 60 試行について 4000 世代の時点で集団の 5% 以上を占めた個体についてその割合を多い順に示したものが図 2.15 である。 $[x00x]GT=[1001]$ が半分弱を占めているのがわかる。この中で $[xxxx]GT=[0001]$, $[xxx1]GT=[0001]$ は、 $[x00x]GT=[1001]$ に次いで高い割合を占めている。この理由は、集団中に頻繁に出現するものの、他の戦略の侵入を容易に許してしまい収束するまでに至らないためである。この点で、完全に収束した $[x00x]GT=[1001]$, $[x001]GT=[1000]$ とは性質が大きく異なる。

また、 $[x001]GT=[1000]$ 型戦略は Boerlijst らが提案した pPAVLOV 戦略[Boerlijst 97] とほぼ同等の振る舞いをするのがわかっている。図 2.16 は両戦略を状態遷移図で表したものである。協調の回復過程において両者は非常に類似しており、どちらも

「2回の裏切りの後に協調し合いに転じる」性質を持っている。今回のモデルのようなシンプルな進化の枠組みにおいて、このような手作りの戦略が現れたのは興味深い。また彼らは、ジレンマゲームにおける戦略において、状況に応じてタグ付けを行う戦略の重要性を主張し、PAVLOV 戦略に代わるロバストな戦略として pPAVLOV 戦略を提案しているが、それ以上にメタ・パブロフ[x00x]型戦略が頻繁に集団中を占めたという結果は、タグ付けの方法の一つとして、メタ・パブロフ学習行列による学習のような強化学習的なタグ付けが有力であることを示していると言える。

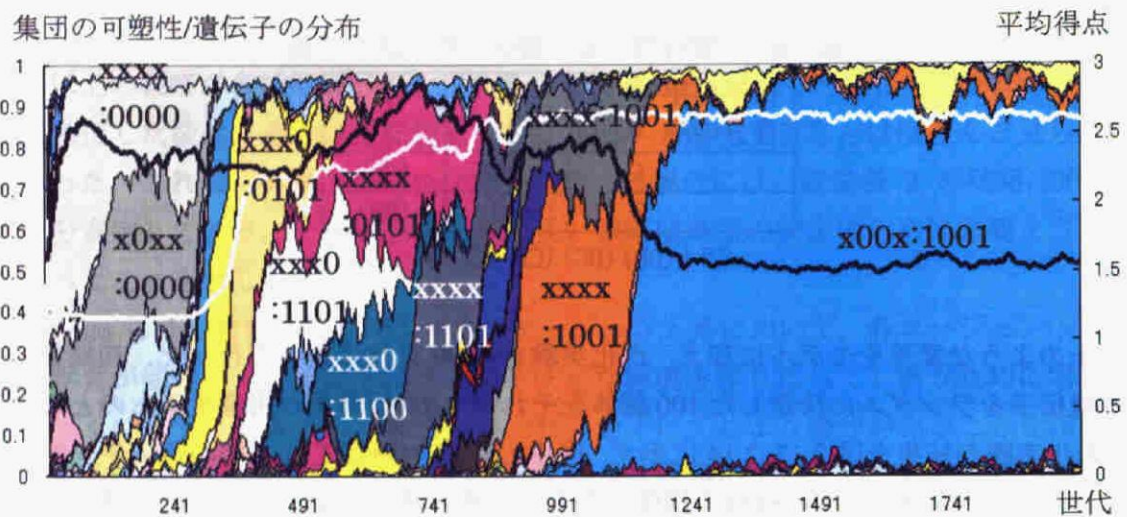


図 2.13: 学習行列を遺伝子に取り込んだ場合の実験結果
(メタ・パブロフ[x00x]型戦略に収束した例)

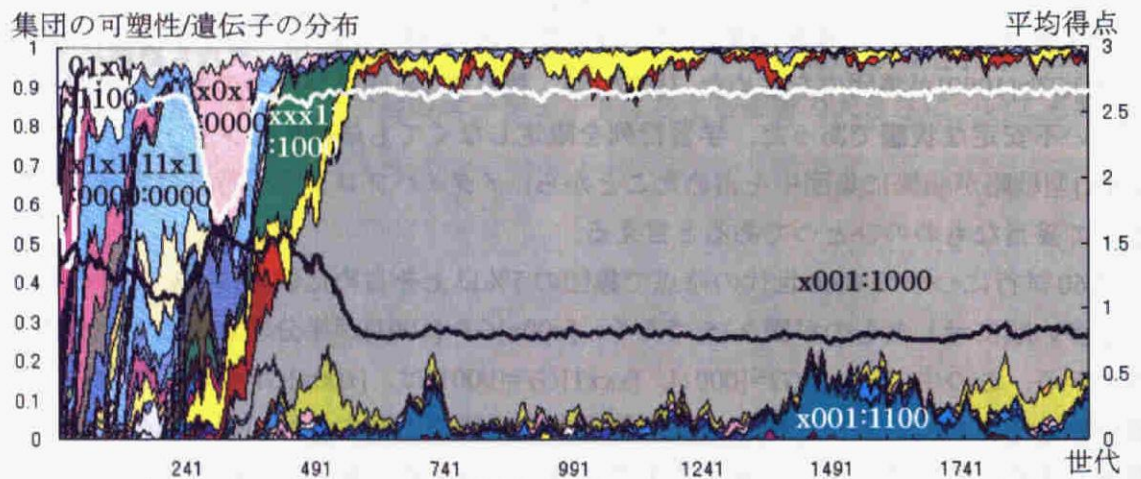


図 2.14: 学習行列を遺伝子に取り込んだ場合の実験結果
([x001], $G^T=[1000]$ に収束した例)

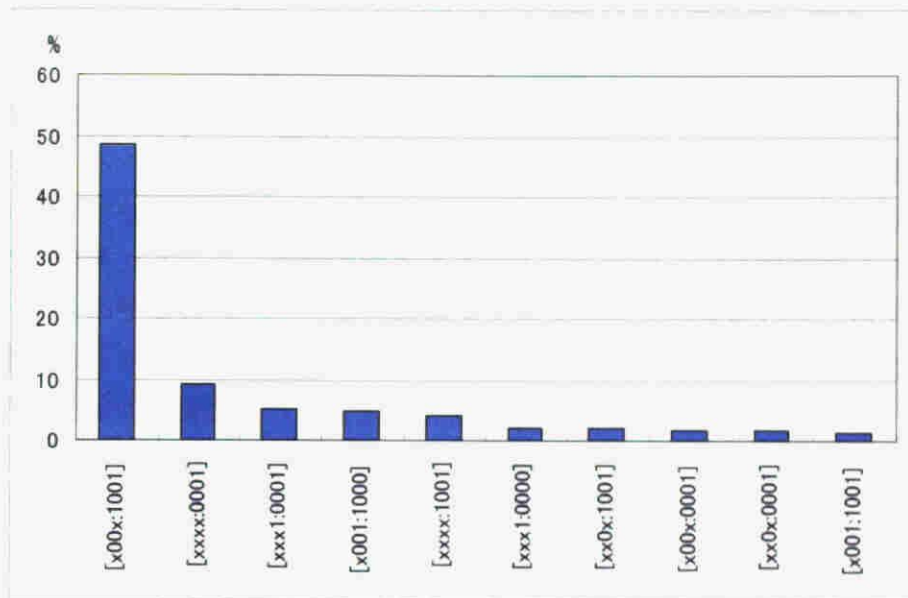


図 2.15: 4000 世代目において各戦略が集団中を占めた割合 (60 試行)

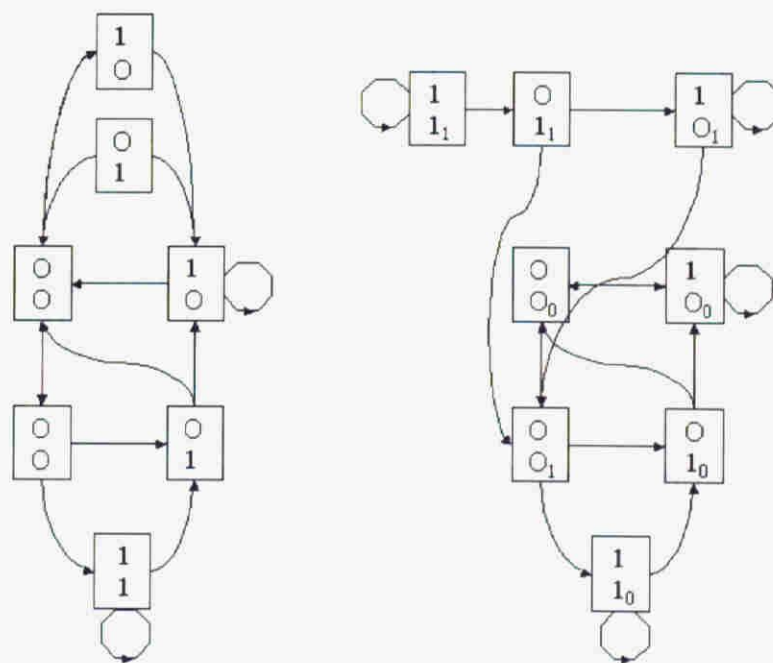


図 2.16: pPAVLOV(左)と[x001], GT=[1000](右)の状態遷

2.6 空間的局所性の影響

これまでの実験では、集団内の個体が他のすべての個体と対戦する総当たり戦を用いて適応度を決定していた。これは、個体間の関係が全く同一であることを意味する。しかし、生態系や人間社会のモデル化、あるいは工学的応用を目的とする場合、現実世界の物理的制約を考慮すると、この条件を緩め、モデルに局所的な相互作用を導入することは本質的であると考えられる。そこで、これまでのモデルに空間的局所性を導入し、表現型可塑性の進化および協調行動の創発に与える影響について解析を行った[Suzuki 00b], [鈴木 02a]。特に本研究では、従来、空間的局所性に関して区別されてこなかった 2 つの要素である、個体間の相互作用に関わる局所性と、世代交代に関わる局所性を明確に区別して拡張を行い、両者の Baldwin 効果に対する影響について議論する。(なお、本論文では、協調行動の創発に対する空間的局所性の影響に焦点を当てた議論を第 3 章で改めて詳細に行う。)

2.6.1 空間的局所性の導入

具体的には、2.3.3 節のモデルに対し次のような変更を行い、空間的局所性を導入した。

- 集団における個体を、 $W \times W$ の 2 次元トーラス状セルに一個体ずつ配置する。(個体数 $N=W^2$ となる。)
- 新たな環境パラメータとして相互作用の規模 i ($1 \leq i \leq (W-1)/2$) を定義し、各個体 (プレイヤー) は図 2.17 に示すように自身を除く近傍 $(2i+1)^2 - 1$ 個体と対戦を行うものとする。
- 次世代に各セルを占める個体を、各セルについて世代交代の規模 r ($1 \leq r \leq (W-1)/2$) によって i と同様にして定義される近傍の $(2r+1)^2$ 個体 (セル上の個体も含む) の中から、対戦で得た平均得点に比例したルーレット選択によって選択する。

新たに定義した 2 つの環境パラメータ i, r の値が小さいほど局所性は高く、大きい

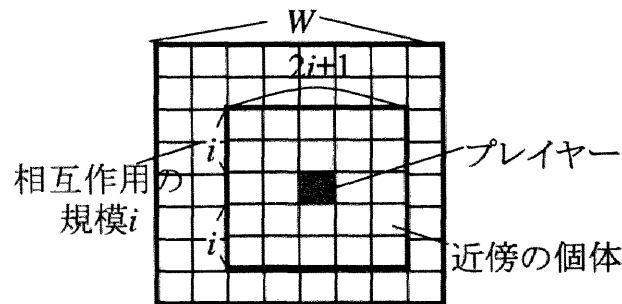


図 2.17: 相互作用の規模 i によって定義される近傍

ほど局所性は低くなる。また、両者とも最大 $i, r=(W-1)/2$ の場合は、局所性を考慮しないこれまでのモデルと同等である。

2.6.2 局所性と可塑性の進化

まず、空間的局所性が与える基本的な特性を把握するため、局所性を表現する 2 つの環境パラメータに同一の値 k を用いて実験を行った。学習規則はメタ・パブロフ学習を用いた。他のパラメータとしては $W=33$, $p_m=1/1500$, $p_n=1/25$, $p_a=99/100$, 世代数 2000 を用いた。図 2.18, 2.19 はそれぞれ、 $k=i=r=1, 2, 3, 5, 10, 16$ を用いた場合における、集団の可塑性と平均得点の推移を表したものである。なお、結果は 100 回の試行での世代ごとの平均である。

まず、全体的な傾向として、これまでと同様 Baldwin 効果の 2 つの段階を通して、最終的に [x00x] 型の戦略が集団中を占めるという進化の過程をたどることが局所性の値 k の大小にかかわらずに共通して観察された。

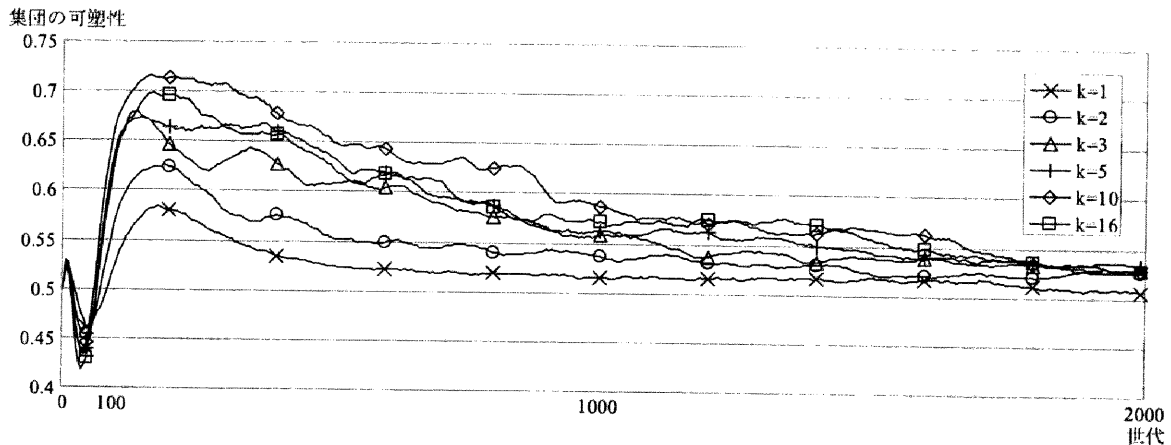


図 2.18：集団の可塑性の推移

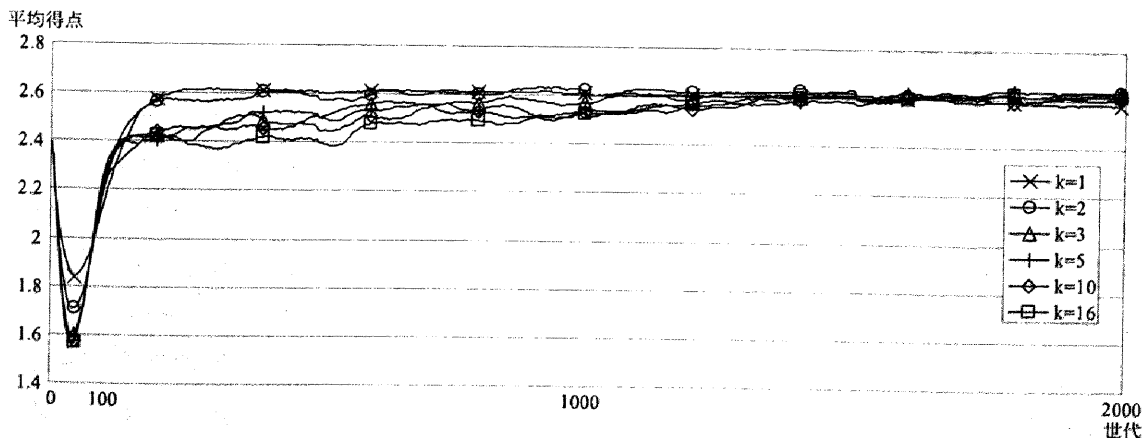


図 2.19：平均得点の推移

しかし、図 2.18 が示すとおり、 k によって集団の可塑性の進化の様子は大きく異なることが判明した。特に、Baldwin 効果の第 1 段階と第 2 段階の境界における可塑性のピークは、 k が小さい場合には低く、 k が大きくなるにつれて高くなる傾向があると言える。このような結果となった要因として、次の点が挙げられる。 k が小さいことは、環境がより局所的であることを意味する。繰り返し囚人のジレンマゲームにおける戦略の進化に空間的局所性を導入した場合には、より協調的な戦略が進化しやすいことが指摘されている[Grim 96]。一般に、この理由は、協調的な戦略は同種同士の対戦で高い得点を得ることができ、協調戦略のクラスを拡大することができることによるとされている。以上は、図 2.19 において、 k が小さいほど、進化の初期段階における平均得点の落ち込みが少ないことから確認できる。このような条件では、進化の初期段階での裏切り戦略の侵入が少ないために、協調集団への進化の際に可塑性は必要とされず、早い段階で[x00x]型の戦略が集団中を占めたと考えられる。逆に、 k が大きい場合には、進化の初期段階において裏切り集団の侵入が顕著であるために、そこから協調集団への進化に十分な可塑性が必要とされたと考えられる。

さらに、図 2.19 において、どの条件でも平均得点は最終的に 2.6 程度に収束するが、 k が高くなるほどその収束が遅い傾向が観察できる。これは、 k が大きくなり、裏切り集団の侵入が容易で不安定な環境となることが、一時的に可塑的でない裏切り戦略が集団中を占める例外的な進化を招いたためであると考えられる。

Kauffman は、自身の提案した NKC モデルを用いて個体間相互作用に基づく動的なシステムの挙動を解析することにより、適応度に関する他の個体との依存度 C が大きいほど、環境は複雑な挙動を示すことを指摘した[Kauffman 93]。本研究において、この依存度は k に相当し、この実験結果は彼の主張を支持すると共に、表現型可塑性の進化は複雑な挙動を示すシステムをより安定させる働きがあることを示している。

2.6.3 相互作用の規模と世代交代の規模

次に、局所性に関する 2 つの環境パラメータを個別に設定した条件で実験を行った。他のパラメータは前節と同様の設定を用いた。図 2.20 は相互作用の規模 i 、世代交代の規模 r を個別に設定して実験を行い、それぞれの条件において、集団の可塑性について試行中で最も高い値を濃淡（黒:0.55, 白:0.8）で示したものである。この可塑性の最高値の高低は、裏切り集団から協調集団への進化の際に、可塑性がどれだけ要求される環境であったかを示している。また、図 2.21 は、試行中で平均得点の最も低い値を濃淡（黒:1.4, 白:1.9）で示したものである。この値が低いほど、進化の初期において裏切りの戦略の侵入が顕著であったことを示している。なお、各値は 5 回の試行の平均である。

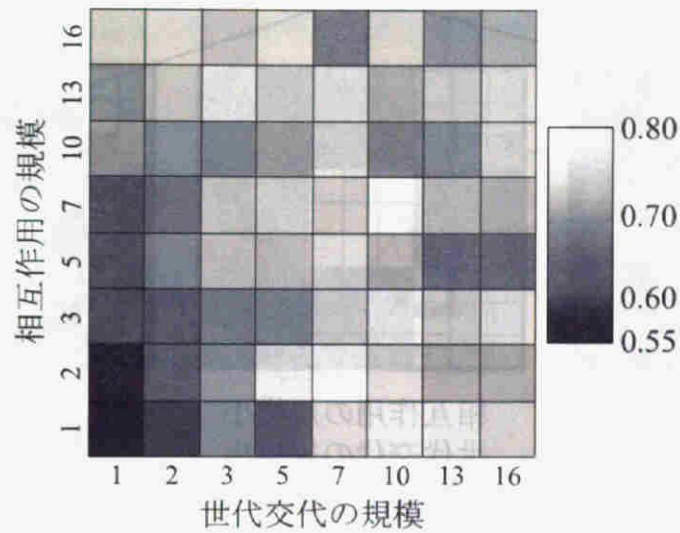


図 2.20：集団の可塑性の最高値

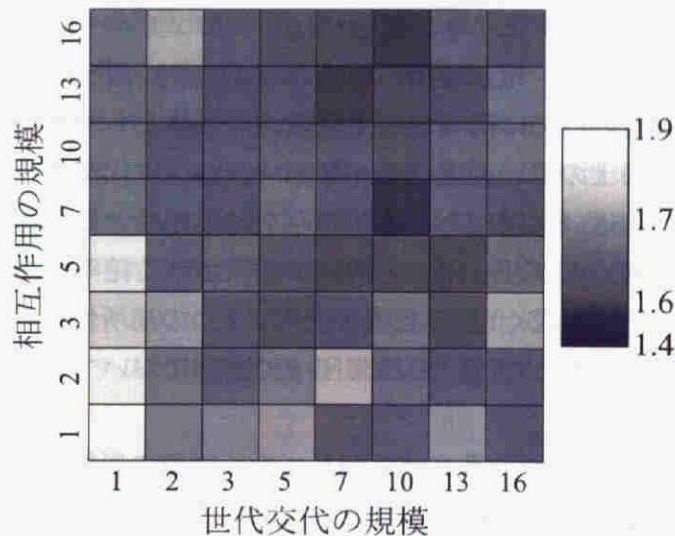


図 2.21：平均得点の最低値

図 2.20 より，相互作用の規模 i ，世代交代の規模 r が共に小さくなるに従って可塑性の最高値は次第に低くなっており，このとき局所性による影響が大きくなっていることがわかる．しかし，一方の局所性のみが高くて，もう一方の局所性が低い状況では局所性の影響がほとんど無いこともわかる．したがって，この環境においては，局所性に関する 2 つの条件が共に満たされた場合にのみ，局所性は進化の過程に大きな影響を与えると言える．以上は，次の理由によると考えられる．前述のとおり，局所性が可塑性の進化に与える影響を考える上で，同種同士の対戦の頻度は，協調戦略が高い得点を得るのに有利であるために重要な要素となる．実際，図 2.21 より，相互作用の規模と世代交代の規模双方が小さい場合において，平均得点の落ち込みは最も小さくなっていることがわかる．

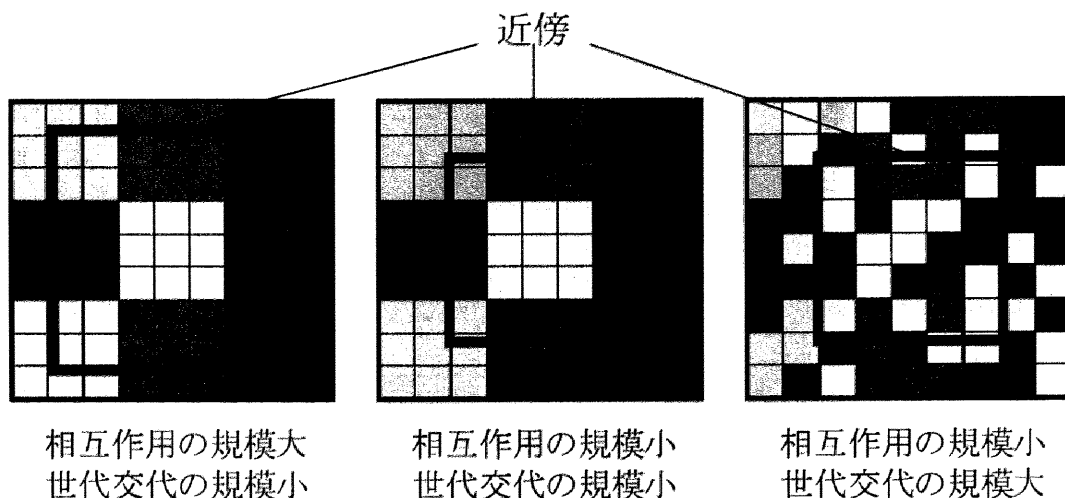


図 2.22：局所性と戦略の分布

ここで、相互作用の規模、世代交代の規模が共に小さい場合と、一方のみが小さい場合の戦略の分布と対戦相手の選択を模式的に図 2.22 に示す。同図より、相互作用の規模、世代交代の規模共に小さい場合には中央のプレイヤーの近傍には同種の個体が多く含まれる。一方、相互作用の規模のみが大きい場合には、戦略の分布は局所的でも近傍が大きいために同種を超えた遠くの個体をも対戦相手として選択することになり、同種同士の対戦頻度は低くなる。また、世代交代の規模のみが大きい場合には、自身から近い距離にある個体のみを対戦相手として選択しても、大域的な世代交代により遠くのグリッドから個体を選択される確率が高いために、その個体が同種である可能性が低くなる。したがって、2つの局所性が高い場合にのみ協調戦略の進化に有利な状況となり、協調集団への進化において可塑性は必要とされなかったと考えられる。

また、2つの局所性は可塑性の最高値に関して全く同等の影響を与えているのではなく、相互作用の規模が小さい場合 ($i=1, 2$) における世代交代の規模の増加が、世代交代の規模が小さい場合 ($r=1, 2$) における相互作用の規模の増加と比べ、可塑性の最高値をやや高くする傾向が見られる。図 2.23 は、 $i=1$ で r を変化させた場合、 $r=1$ で i を変化させた場合の可塑性の最高値の推移を示したものである。同図より、 $i=r=1$ の設定時からの i および r の増加に伴い、ばらつきはあるものの r の増加の方が高い可塑性の最高値をもたらしている（世代交代に関する局所性の減少の方が、可塑性の最高値の増加の抑制をより弱めている）ことがわかる。これは、世代交代の規模が大きいことによる全域的な世代交代が、適応的な戦略が集団全体に急速に伝播するのを促進するために、特に進化の初期段階において適応的な裏切り戦略の侵入を容易にしているためであると考えられる。

これら 2 つの局所性の設定を変更することは、戦略個体の側から見れば、自身の置かれる環境の特性が変化することを意味する。この実験結果は、環境の違いに対

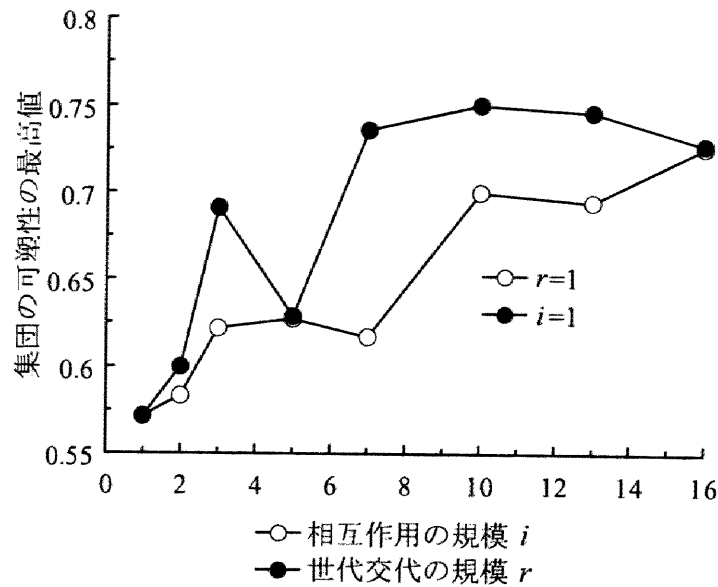


図 2.23 : 2つの局所性の変化に対する可塑性の最高値の比較

して可塑性進化のメカニズムが柔軟に対応し、協調集団への進化を導いたことを示している。現実世界において、局所性が集団全体の挙動に大きな影響を与えていることは自明である。しかし、現実世界が実際どの程度の局所性の下に成り立っているかを知ることは容易ではない。これらの結果は、どのような局所性を持った環境においても、表現型可塑性の進化がその影響を柔軟に吸収し、協調集団への進化を導いたことを示している。これは、表現型可塑性の進化の枠組みが、より現実的なマルチエージェントシステムの進化的獲得において重要な要素となることを示唆するものである。

2.7 まとめ

本章では、動的な環境における進化と学習の相互作用について知見を得るために、繰り返し囚人のジレンマゲームにおける戦略にメタ・パブプロフ学習に基づく表現型の可塑性を導入して進化実験を行い、その結果について考察してきた。

初めに、基本的な実験として記憶長 2 の戦略集団を用いて実験を行った。ランダム型学習を用いる集団では、ある特定の戦略に完全に収束することはなかったが、メタ・パブプロフ学習を行う集団においては、Baldwin 効果の 2 つの段階が働いた結果 [x00x]型の戦略が集団中を占めた。

Baldwin 効果が確認された過程において、可塑性が戦略の進化に与える影響を具体的に解析した。その結果、集団は直接協調的で安定した集団へと進化するのではなく、Baldwin 効果の第 1 段階として一旦集団の可塑性が増加する方向へ進化し十分な協調関係を実現してから、第 2 段階として可塑性の減少とともに協調的で安定した集団へと進化した。本研究においては学習の際の明示的なコスト（およびメリット）を与えないとしたが、解析の結果、このモデルにおける学習のメリットは裏切りのな集団から協調的な集団へ進化するために必要とされ、またコストは他の戦略に侵略される可能性の大きさとして働いていることがわかった。このメリットとコストは、どちらも動的環境に特有のものであり、動的環境における Baldwin 効果の存在を示すものである。

最終的に集団の大部分を占めたメタ・パブプロフ型[x00x]戦略の解析を行ったところ、[x00x]型戦略は記憶長 2 の集団において ESS の条件を満たすなど、必要最低限の可塑性を持った強力な戦略であることが判明した。また、記憶長の突然変異や学習行列の進化を導入し、より一般化したモデルにおいても、多くの場合において[x00x]型戦略が集団中を占めることがわかった。これは[x00x]型の戦略が純粋にジレンマゲームにおける強力な戦略として興味深いだけでなく、このような進化と学習の割合が自動的に調節される進化の枠組みが、動的な環境に対する集団の安定性を実現するための重要な要素となっていることを示唆する点で意義深いと考えられる。

最後に、これまでのモデルに相互作用・世代交代に関する空間的局所性を導入し、Baldwin 効果への影響について解析を行った。その結果、空間的局所性の導入により、導入前と比較して、Baldwin 効果の第 1 段階における可塑的形質の頻度の増加が強く抑制されることが明らかになった。これは、局所性の導入が、協調集団への進化に有利な環境を作り出し、進化の初期段階における裏切り戦略の侵入を減少させるために、協調集団への進化に可塑的形質が必要とされなかったことによると考えられる。また、相互作用に関する局所性は対戦相手の候補を制限し、世代交代に関する局所性は戦略の分布を制限することにより、進化の過程に異なった形で影響するが、どちらも Baldwin 効果における可塑的形質の頻度の増加を抑制する傾向があること

が明らかになった。しかし、その影響は、片方の局所性のみが高い場合では小さく、両者が高い場合において最も大きくなり、このとき可塑的形質の頻度の増加が最も強く抑制されることが判明した。また、両者の影響の大きさを比較した場合、世代交代に関する局所性の方が、可塑的形質の頻度の増加をより大きく抑制する傾向があることなども明らかになった。

本研究で示唆される、生物における2つの適応プロセスが形作るイメージは、「学習が進化の方向性を見つけ、次に、進化をその方向に引き寄せる」というものである[有田 00]。これは、従来の「ランダムな突然変異の繰り返しにおいて自然選択が方向性を決めて進化が進んでゆく」というオーソドックスなものとは異なる。この過程を、生物の適応プロセスの理解をサポートする概念である、適応度地形を用いると、次のように表現することができる。

まず、図 2.24 に示すような、遺伝子型の空間と、その遺伝子型に対応する表現型の適応度を示す空間（適応度地形）を考える。個体、または種がこの遺伝子空間上の一点で表されるとした場合、従来の自然選択による進化は、突然変異によって遺伝子空間の近傍をランダムに探索し、適応度の高い表現型に対応する遺伝子型のほうへ動いていくというものであった。一方、今回のイメージでは、個体、あるいは種全体を「尺取虫」に比喻することが可能である。尺取虫の頭を表現型空間、尾を遺伝子型空間に置く。まず、表現型空間上でより適応度の高いほうへと頭を伸ばす（第1段階）。このとき、尾を動かさずに頭を何回か別の方向に動かすこともできる。そして、尺取虫の体長が最も長くなる（可塑性最大）と、今度は尾を頭のほうに動かす（第2段階）。このようにして、先行する頭（学習）の行く先によって、尾（進化）の道筋が方向付けられるというシナリオである。今回の実験では、外的環境（利得行列の値など）を固定したが、これを変動させればこのような伸縮が繰り返されていくと考えられる。

本研究はジレンマゲームにおける戦略の進化という抽象的な環境での議論である

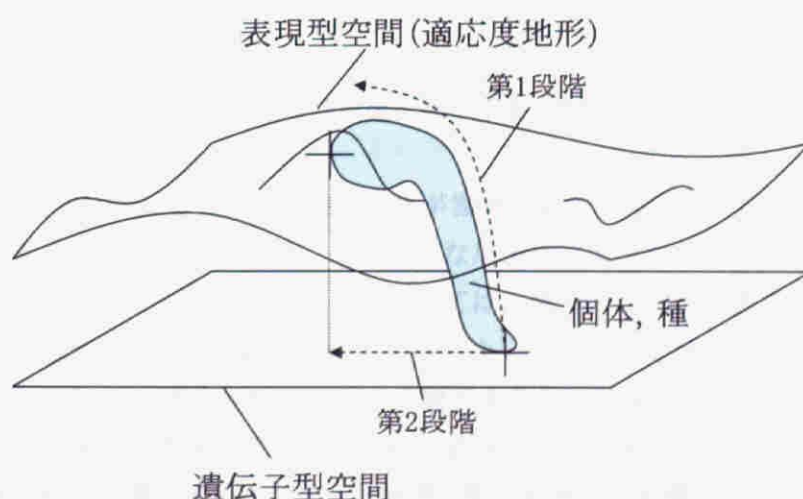


図 2.24：進化と学習の2つの適応プロセスが形作るイメージ

が、それゆえに多様な発展が可能であると考えられる。その可能性としていくつかの方向性を挙げる事が出来る。ひとつは、現在のモデルについての解析および拡張を更に行い、進化と学習の相互作用に関する考察を深めることである。たとえば、ジレンマゲームの利得行列を変動させて、環境自体の変動を導入するのは興味深い。

もうひとつは、このような動的環境における進化と学習の相互作用の仕組みの工学的応用である。本研究で扱ったメタ・パブプロフ学習を最もシンプルな強化学習メカニズムの一つと見なすと、集団全体はマルチエージェント強化学習系として捉えることができる。近年、強化学習に関する研究が盛んに行われている中で、マルチエージェント系における強化学習では、個々のエージェントの学習が互いの環境を不安定にし、学習を困難にさせてしまうことは一つの問題である。このような問題に対し、本研究で用いたような、集団全体の学習への依存度が量的に定義され、かつ、依存度が進化の過程で調節されるようなモデリングは有効であると考えられる。また、このような手法は、進化と学習の相互作用のみならず、遺伝子とミームの共進化といった、複数の適応プロセスが一つの系に存在するような一般的状況において適用可能であり、興味深い現象をもたらすことが期待される。

また、より生物学的な側面への発展のひとつとして、鳥の鳴き声の性選択による進化に関する研究が挙げられる[Todd 96]。鳥の鳴き声は、幼少期における成鳥からの学習に依存する要素と、先天的に決定されている要素に分けられる。その比重は種や環境によってさまざまであり、これは進化の過程でうまく調節された結果であると推測できる。この過程において、本研究で示したような進化と学習の相互作用が起こっている可能性は十分考えられる。モデルを構築し、様々な環境で実験を行うことで、鳥の鳴き声の進化に関する新たな知見が得られることが期待できる。

第3章 協調行動の創発に対する空間的局所性の影響とその進化

利己的集団における協調行動の進化に関して、繰り返し囚人のジレンマゲームの戦略の進化を用いた研究が多くなされてきた。2章では、ゲームを行う戦略集団を、動的要因を集団の構成要素自体が内包する環境として採用し、動的環境における進化と学習の相互作用について様々な議論を行ってきた。繰り返し囚人のジレンマゲームが極めて単純なゲームであるにもかかわらずこのような興味深い議論が可能である理由の一つは、社会に生じるジレンマ状況を的確に抽出したものとなっていることが挙げられる。

一方で、より現実的・一般的な設定での議論を目指して、繰り返し囚人のジレンマゲームにおける戦略の進化に対していくつかの拡張を行ったモデルを採用した研究も同時になされてきた。その一つに、空間的局所性の導入が挙げられる。これは、現実の社会構造を考えると、各個体が他のすべての個体と直接関係しているのではなく、ある程度の局所性を持って相互作用を行っていると考えられるためである。しかし、従来の空間的局所性を導入した従来の研究は、空間的局所性に関して区別して考慮することのできる2つの要素、すなわち個体がゲームを行う対戦相手を選択する際に生じる局所性と、各空間的位置における次世代の個体を選択する際に生じる局所性を同一のものとして扱っている点で、局所性の影響を十分議論していないと考えられる。

そこで、本章では、空間的局所性に関して本質的に異なる要素である、相互作用に関する局所性と世代交代に関する局所性について、それぞれが協調行動の創発にどのような影響を与えるかについて知見を得ることを第一の目的とする。このために、ジレンマゲームのもう一つの拡張であるN人版繰り返し囚人のジレンマゲームの戦略を1次元セル上に配置し、各個体がゲームの戦略の他にどれくらい近傍の個体と相互作用・世代交代を行うかという相互作用の規模・世代交代の規模が個別に定義できる進化モデルを構築し、それぞれが協調行動の創発と集団の進化ダイナミクスに与える影響について解析を行う。(なお、2.6節での相互作用の規模、世代交代の規模に関する議論は同様の視点に基づく研究の一つである。)

本研究の第二の目的は、2つの局所性の進化が協調行動の創発に与える影響について知見を得ることである。従来の研究では、このような局所性は環境から一意に与えられる要素として扱われてきた。しかし、現実世界においては、個体それぞれが独自の相互作用や世代交代に関する戦略を持ち、それが進化してきたというシナリオは十分考えうるものである。そこで、これら2つの局所性を決定する要素を各個体が遺伝的なパラメータとして保持し、それに従って相互作用・世代交代を行う状況を上記のモデルにおいて設定し実験を行い、各パラメータの進化ダイナミクスと協調行動の創発への影響について解析を行う。

本章の構成は次のとおりである。3.1節では、本研究の背景と、本研究における空間的局

所性に対するアプローチを明確にした上で、目的について述べる。3.2節では、構築したモデルについて解説する。3.3節では、2つの局所性が協調行動の創発に与える基本的影響について、それぞれの進化なしの場合を用いて明らかにする。3.4節では、局所性の進化を導入し、両者の進化ダイナミクスについて議論する。3.5節で、これらの結果をまとめ、今後の課題について述べる。

3.1 背景と目的

繰り返し囚人のジレンマゲームは、社会的集団に生じるジレンマ状態をシンプルに表現したものであり、これまで数多くの研究がなされてきた。(繰り返し囚人のジレンマゲームに関する詳細な解説は 2.1.1 節を参照されたい)。一方、ジレンマゲームは、そのシンプルさゆえに、現実世界と比較すると極端に単純化された点があるため、より現実的なモデル化のためにいくつかの拡張がなされてきた。その代表的な例が空間的局所性の導入や、N 人版ゲームへの拡張である。本節では、これらの拡張を導入した先行研究を挙げた上で、これらの拡張に対する本研究でのアプローチについて明確にし、研究の目的について述べる。

3.1.1 空間的局所性と協調行動の創発

繰り返し囚人のジレンマゲームのようなゲーム論的状況設定に基づく進化モデルにおいては、集団内のすべての個体が他のすべての個体と対戦を行い、適応度を決定する総当たり戦を用いることが一般的である。しかし、現実世界は必ずしも個体間の相互作用の関係が均一ではないため、現実的、生物学的により妥当な設定として空間的な局所性を導入した議論がなされてきた(図 3.1)。

一般的には、空間的局所性の導入は集団に多様性をもたらすことが知られている。Nowak らは、2次元セル上に「常に裏切り」または「常に協調」という単純な2つの決定論的戦略を配置した[Nowak 92a]。各個体は、周囲8近傍の個体のみとそれぞれ囚人のジレンマゲームを行った結果を適応度とし、各セルの近傍9個体のうち最も適応度の高い個体が、次世代のセルを占めるという進化メカニズムを採用し、実験を行った。その結果、裏切りに対する誘惑(メリット)を示すゲームの利得行列の値が中程度であるとき、2つの戦略が空間上を遷移する、カオス的なパターンが多様に現れることが判明した。また、Lindgren らは、2.5.1 節で採用した戦略遺伝子列と同様の、記憶長の突然変異を含むオープンエンドな戦略の進化を用いて、ゲームを定義するパラメータが戦略の空間的パターンの分布に対して与える影響について議論している[Lindgren 94]。このとき、戦略個体を2次元平面上に配置し、各個体は

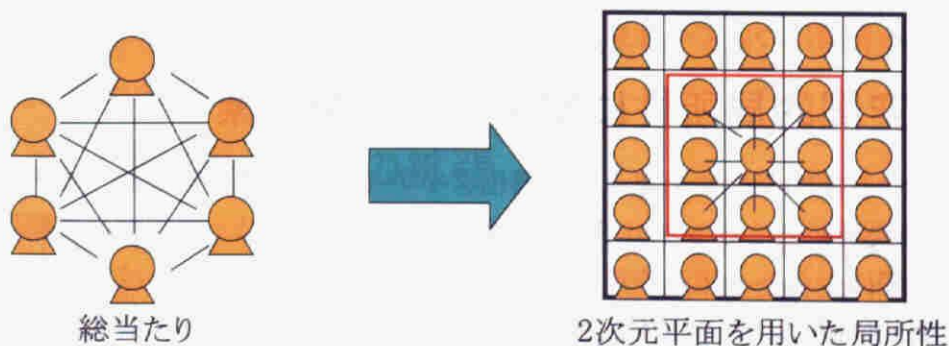


図 3.1 : 空間的局所性の導入

周囲 4 近傍の個体と対戦を行い、次世代の各セルを占める個体はそのセルの周囲 4 近傍内で最も成績の良い個体とする進化メカニズムを採用した。彼らは、すべての個体が他のすべての個体と対戦を行う総当たり戦での進化のシナリオと、空間的局所性を導入した場合のシナリオを比較し、例えば記憶長 1 の戦略群では ALLD (全面裏切り) 戦略, ALLC (全面協調戦略) と ATFT (しっぺ返し戦略の逆の手を常にとる) 戦略が共存するなど、局所性を導入したモデルの方がより多様な戦略が混在しながら記憶長が進化することを示した。

また、空間的局所性を導入した繰り返し囚人のジレンマゲームにおける戦略の進化においては、局所性のない場合と比較して協調行動が築かれやすいことが指摘されている。これは、協調的な戦略は、互いに協調し合うことによって適応度を増加させ、空間上にクラスタを作り拡大することができるのに対し、裏切りの戦略は同種同士で裏切り合って適応度を下げてしまい、クラスタを維持しにくいためであるとされている。

Grim は、2 次元平面上にジレンマゲームの戦略個体を配置し、各個体は平面上の近傍のみゲームを行うという設定を用いて進化実験を行った[Grim 96]。彼は、各個体の戦略として、繰り返し対戦の初回に出す手と、一つ前の対戦において相手が協調・裏切りを出したそれぞれの場合において、今回の対戦で協調する確率を遺伝的に記述し、戦略を進化させた。このとき、局所性のない環境と比べ、相手の裏切り行為をある程度許容するという意味で寛大なしっぺ返し戦略として知られる“GTFT” (初回は協調する。それ以降は確率 1/3 で協調し、それ以外はしっぺ返し戦略をとる) [Nowak 92b]よりも 2 倍協調的である個体 (相手が前回裏切りを選択したときに次回協調する確率が 2/3) が集団を支配可能であることを示した。

一方、田村は、Grim のような 2 次元平面による空間的局所性を導入しながらも、しっぺ返し戦略や Pavlov 戦略を含む十数種のあらかじめ定義された戦略群を用いて進化実験を行った結果、Friedman と呼ばれる、初回から協調し続けるが相手が一度裏切るとそれ以降は常に裏切りつづける戦略[Axelrod 84]が集団中の多くを占めることを指摘した[田村 02]。この戦略は、相手の裏切り行為を最後まで許さない (その後協調に転じない) という点で、寛大でない戦略と言え、Grim らの主張とは異なる結論と言える。このことは、モデルにおいて記述可能な戦略の種類が進化ダイナミクスに与える影響についても注意を要することを示唆する点で重要であると言える。

3.1.2 空間的局所性に関する 2 つの要素

このような空間的局所性を導入した進化ダイナミクスを考える上で、局所性は大きく 2 つの要素に分けることができると考えられる。一つは、集団内の各個体がゲームを行う際に生じる、相互作用に関する局所性である。もう一つは、次世代において各空間的位置を占める個体を選択する際に生じる、世代交代に関する局所性で

ある。(厳密に言えば、選択も適応度を介した個体間の相互作用によってもたらされると見なすことも可能ではあるが、本研究では、選択をゲームを行う際の相互作用と明確に区別して、そこに生じる局所性を世代交代に関する局所性と表現する。)

例えば、前述のような2次元平面上での戦略進化を考えると、各個体がどれくらい近くの個体とゲームを行うかを決定する近傍の領域の大きさが、相互作用に関する局所性を表すパラメータとなりうる。また、世代交代に関する局所性については、次世代において各セルを占める個体を選択する際、注目するセルからどれくらい近傍の個体群を候補とするかが、世代交代に関する局所性を表すパラメータとなりうる。これら2つの要素の進化ダイナミクスに与える影響が本質的に異なることは容易に想像できるにも関わらず、それぞれが協調行動の創発に与える影響については、これまであまり議論されてこなかったトピックの一つである。

3.1.3 空間的局所性の進化

また、これまで行われてきたほとんどすべての研究において、このような局所性に関するパラメータは環境から一意に与えられるものとして定義されてきた。しかし、現実社会を考えると、例えば、あまり付き合いの良くない人は隣人との付き合いを避けたり、社交的な人はより遠く、たくさんの人と面識を持ったりなど、各個体によって近傍の個体との付き合いの仕方が異なるという側面がある。つまり、集団ははじめからある同じ規模で相互作用を行っているのではなく、各個体のそれぞれの特性としてどれくらいの規模で相互作用を行うかが決定され、それに従って実際の相互作用が行われるということである。

ここでもし、このような相互作用に関する各個体の特性が部分的にでも遺伝的に決定される状況を仮定すると、進化の過程において他の個体とうまくやっていけるような特性へと進化するというシナリオが考えられる。従って、相互作用の規模が、環境パラメータというよりもむしろ各個体によって戦略の一部として保持されるといった、相互作用に関する局所性の進化が導入された状況も議論すべき重要な問題設定であると言える。

世代交代の規模に関しても、自身がどれくらいの範囲において生存競争に参加するかは個体の生存にとって極めて重要な要素であり、これについても戦略の一部として遺伝的に保持するという設定は十分考えられうるものであり、2つの局所性そのものの進化に関する議論は重要であると考えられる。

3.1.4 N人版ゲームへの拡張

ところで、繰り返し囚人のジレンマゲームにおける戦略進化と協調行動の創発に関するモデルでは、これまで述べてきたように個体が置かれる環境に対して拡張が行われてきただけでなく、行うゲーム自体に対してもより現実的な設定を目指して

拡張が行われてきた。

その一つが2人版ゲームのN人版ゲームへの拡張である。現実世界における意思決定は、一般的には2人以上の構成員によって行われるという観点から、2人版繰り返し囚人のジレンマゲームのエッセンスを維持しつつも3人以上の構成員で行うことができるように拡張した、N人版繰り返し囚人のジレンマゲームに関する研究がなされている。一般に、N人版のゲームは次の手順で行われる[Yao 99].

- ゲームの構成員数を n とする。
- 各プレイヤーは、他のプレイヤーの手を知らないまま、同時に協調 (C) または裏切り (D) を同時に選択する。
- n 人中の協調者の人数 c に応じて、協調者は利得式 $C(c)$ 、裏切り者は利得式 $D(c)$ によって表される得点を得る。
- 以上の対戦を繰り返し行う。

このとき、ジレンマ状況を表すために、利得式は一般に次の2つの条件を満たす必要がある。

$$\text{(条件 1)} \quad D(c-1) > C(c) \quad \text{for } 1 \leq c \leq n \quad (3.1)$$

$$\text{(条件 2)} \quad C(n) > D(0) \quad (3.2)$$

条件 1 は、他のプレイヤーが次回も同じ手を維持するならば、自分は裏切ったほうが高い利得を得ること、条件 2 は、もし全員が裏切ると、全員が協調するよりも全員にとって悪い結果を招いてしまうことを示す。これら2つの条件によって、2人版と同様のジレンマ構造を抱えた状況設定を表現すると考えることができる。

例えば、Joshi は、2人版繰り返し囚人のジレンマゲームにおける利得行列の各値を、表 3.1 に示すとおり R, T, S, P として、協調者の利得 $C(c) = [(c-1)R + (n-c)S]/(n-1)$ 、裏切り者の利得 $D(c) = [cT + (n-1-c)P]/(n-1)$ と定義した (表記の方法は本研究での表記法に改めてある) [Joshi 87].

表 3.1 : 囚人のジレンマゲームの利得行列

相手の手 (→)	協調	裏切り
自分の手 (↓)	(C)	(D)
協調 (C)	(R, R)	(S, T)
裏切り (D)	(T, S)	(P, P)

$$2R > T + S$$

(自分の得点, 相手の得点)

これらの式によって得られる利得は、各プレイヤーがあらかじめ D または C の手を決めておき、その手に従って他のすべてのプレイヤーと2人ゲームを行ったときに得られる利得の平均に等しく、2人ゲームの利得行列を考慮した定義であると言える。彼は、この利得式を用いて、しつぺ返し戦略と全面裏切り戦略との共存可能性に対する、 R, T 、繰り返し対戦回数 M 、集団のサイズ N の影響について数理的解析を行った。その結果、 R, M が小さくなるほど、また、 T, N が大きくなるほど、し

っぺ返し戦略が全面裏切り戦略と共存するために最低限必要な集団中の比率が大きくなることが明らかになった。

また、Matsushimaらは、3人版繰り返し囚人のジレンマゲームをJoshiらと同じ利得式で定義した上で、他の2個体の一つ前の手に依存した決定論的戦略を有限状態遷移機械として表現して進化実験を行っている[Matsushima 98]。彼らは計算量の軽減を考慮し、個体を2次元平面上に配置し、各個体は周囲9近傍のうち自身を含む可能な12の組み合わせにおいて3人版の繰り返し囚人のジレンマゲームを行い、周囲9近傍で最も成績の良い個体が次世代の各セルを占めるという進化メカニズムを採用した。その上で、ある戦略で密集した空間の中への新たな戦略の侵入可能性に注目し、侵入する側とされる側にそれぞれ置かれた場合に用いられる遷移規則の構造を分離することで、進化の過程で出現した戦略の特性について論じている。実験の結果、例えば、ALLC型（全面協調）とFriedman型（はじめから協調し続けるが、一度裏切られると常に裏切り続ける）といった特徴が組み合わさった戦略が出現し、他種の侵入や自身の突然変異種の出現に対してロバストであったと報告している。

この研究で興味深いのは、ノイズ有りの2人版ゲームと、ノイズなしの3人版ゲームとの進化ダイナミクスの類似性について言及している点である。彼らの主張では、3人ゲームは、3人のうちある2人がゲームを行っていて、残りの一人がノイズの発生源となっていると見なすことができるとしている。実際、進化実験の結果では、Lindgrenらが行ったノイズありの2人ゲームにおける決定論的戦略のオープンエンドな進化のシナリオ[Lindgren 91]に登場する戦略の特徴を備えた多彩な戦略が登場した。もちろん、ノイズなしの3人ゲームがノイズ有りの2人ゲームと全く同値であることは考えにくいだが、ゲームを行う人数を増加させることが、ノイズを導入した場合と同様に系全体に多様な進化をもたらしたという点で興味深いと言える。

一方、Bankesは、 $D(c)=c$ 、 $C(c)=c-K$ (K は定数) という、より単純な利得式を採用した[Bankes 94]。Joshiの利得式と異なる点は、Joshiの利得式は表3.1のパラメータの組み合わせにより、集団内の協調者の数が増えるに従って協調者と裏切り者の利得の差が大きくなったり小さくなったりする場合があるのに対し、Bankesの利得式では常に一定であることである。(なお、Joshiの利得式において、表2.1のような一般的な2人版ジレンマゲームにおける利得の行列を用いた場合には、協調者と裏切り者の利得の差は協調者の増加とともに小さくなる。また、互いに協調した場合に得られる利得のみ $R=4.0$ に変更するとBankesの利得式と同様、常に一定となる。)

彼は、 N 人ゲームの戦略を文化的自己複製子であるミームとして表現し、プレイヤーごとに用意されたミームのプールが、繰り返し対戦の成績に応じたルーレット選択によって独立に進化するモデルを用いた。協調することが全員裏切るよりもよい結果となる最小の構成員数を決めるパラメータ K と協調行動の創発との関係について解析を行い、協調行動の進化にとって、一般的に言われるゲームの構成員の数

よりも、むしろ注目するパラメータ K が重要であることを指摘した。このように、 N 人ゲームへの拡張により、集団は従来の 2 人版繰り返し囚人のジレンマゲームとは異なった進化の様相を呈すると言える。

3.1.5 目的

以上を踏まえ、本研究は、これまで同一のものとして議論されてきた 2 つの局所性のそれぞれが協調行動の創発にもたらす影響、および、局所性の進化を導入した場合における協調行動の創発への影響と局所性の進化のダイナミクスに関して知見を得ることを目的とする [鈴木 01a, 01b, 01c], [Suzuki 01, accepted a]. なお、本研究で想定する空間的局所性は、単に地理的な隔たりの抽象化に限らず、個体間の付き合いの関係といった、集団内での相互作用に関わる局所性一般を抽象化するものである。

具体的には、繰り返し囚人のジレンマゲームにおける協調行動の創発に対する空間的局所性に関し、次のような問いに答えることを目的とする。

- 相互作用に関する局所性と世代交代に関する局所性について、それぞれが環境パラメータとして一意に与えられる場合、協調行動の創発にどのような影響を与えるだろうか？また、このとき N 人ゲームの構造はどのような影響をもたらすだろうか？
- 2 つの局所性が環境パラメータではなく、各個体の戦略の一部として進化する場合、その進化ダイナミクスはどのようなもの？また、それは協調行動の創発を促進するだろうか？協調行動の創発をもたらすとすれば、それはどのような条件においてであろうか？

このために、本研究では、 N 人版繰り返し囚人のジレンマゲームの戦略を 1 次元セル上に配置し、各個体がゲームの戦略の他にどれくらい近傍の個体と相互作用・世代交代を行うかという相互作用の規模・世代交代の規模を遺伝子として保持する進化モデルを構築した。 N 人版ゲームを採用する理由は、これまでに述べた相互作用の規模の影響を考える場合には、単にグループ内で 2 人版のジレンマゲームを複数回行うよりも、そのグループ全員で N 人版ゲームを行う方がより現実的であるだけでなく、集団内で複数のグループが構成される場合には、その大きさがそれぞれ異なることによる影響が期待されるためである（この点については、3.2.1 節で改めて述べる）。

石渕らは、空間型繰り返し囚人のジレンマゲームにおいて、これら 2 つの局所性が協調集団への進化に与える影響について解析している。彼らはノイズ率が協調集団への進化に与える影響に注目し、ノイズ率が中程度の場合において、前述の 2 次元平面における相互作用の規模、世代交代の規模に関して、どちらも大きくなるほど協調集団への進化が難しくなることを示している [Ishibuchi 99], [石渕 00]。また、

本研究では、2.6 節において、表現型の可塑性を導入した繰り返し囚人のジレンマゲームの戦略の進化モデルを 2 次元平面に拡張し、2つの局所性の度合いに応じて、戦略進化の過程に与える表現型可塑性進化の影響が異なることを示した。本研究では、N 人版ゲームを用いた状況設定を採用することにより、協調行動の創発に対する 2つの局所性の影響に関して、これらとは異なる知見を得ることが期待できる。

3.2 空間的局所性を導入したN人版繰り返し囚人のジレンマゲームにおける戦略の進化モデル

以上を踏まえ、N人版繰り返し囚人のジレンマゲームの戦略を1次元セル上に配置し、各個体がゲームの戦略に加えどれくらい近傍の個体と相互作用・世代交代を行うかという相互作用の規模・世代交代の規模を遺伝子として保持する進化モデルを構築した。本節では構築したモデルについて詳細に解説する。はじめに、本研究で採用するN人版繰り返し囚人のジレンマゲームについて解説し、次に、空間的局所性を導入した進化メカニズムについて、2つの局所性によって形成されるグループに注目して解説する。

3.2.1 N人版繰り返し囚人のジレンマゲーム

本研究では、Banksの利得式を元にした次のような利得式を採用する[Banks 94] (図3.2)。

$$\text{協調者の利得} \quad C(c) = c/n \quad (3.3)$$

$$\text{裏切り者の利得} \quad D(c) = (c+k)/n \quad (3.4)$$

ここで、 n はゲームを行うグループの大きさ、 k は協調するためにかかるコストを表す定数 ($1 < k < n$) とする。Banksの利得式と異なる特徴は、

特徴1) 利得の値全体が集団のサイズ n で正規化されている。

特徴2) 協調するためにかかるコスト k が裏切り者の利得の方に上乘せされている (裏切ることによるメリットと見なすこともできる)。

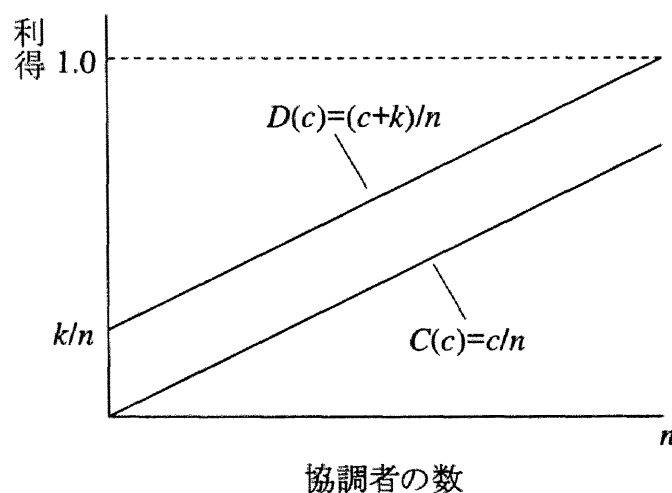


図3.2 : N人版ゲームの利得式

ということである。特徴1)は、本研究で採用する進化モデルでは一つの集団で複数の構成員数が異なるグループをつくって対戦を行うため、どんな構成員数のグループでも全員が協調したときに得られる利得を1.0に統一することによって、単に構成員数が大きいことによる直接的なメリットをなくすためのものである。特徴2)は、対戦の平均利得を適応度にする際に、適応度が負の値をとらないようにするためのものである。また、この設定は具体的には次のような状況が当てはまると考えられる(図3.3)。

n 人で共有財を利用する。財を最大限に利用するための維持費用としてコスト k が利用人数に関わらず必要となる。そこで、 k を構成員数で均等に配分した k/n をそれぞれが負担しようと約束するが、実際に各構成員がコストを支払う(協調する)か支払わないか(裏切る)は個人の自由である。ただし、各構成員が共有財から得られるメリット(財の分配)は実際に支払われたコストの総額に比例する。

Bonacichらは、構成員の増加が協調行動の創発を促進するかどうかは、構成員数の増加に伴う利得式の変化に依存すると指摘した[Bonacich 76]。彼らは、利得式から計算可能な、協調することのメリット($C(n)-D(0)$)と裏切りへの誘惑($D(i-1)-C(i)$, $i \in [1, n]$)が、集団の構成員数の増加に伴ってどのように変化するかによって利得式を5種類に分類している。本研究で採用する利得式は、先に述べた特徴1)により集団の構成員数 n で正規化されているために、構成員数が増加するほど協調者のメリットは増加し、裏切りへの誘惑は減少するタイプに属する。このタイプに属する利得式では、構成員の増加に伴い協調行動がより成立しやすくなる可能性がある。前述のとおり、このような状況設定も現実的であるにもかかわらず、これを用いた議論はあまりなされていないため、彼らはこのような条件でのN人版ゲームの議論の重要性を主張している。また、本モデルでは、一つの集団に複数の異なる大きさのゲームを行うグループが存在可能である。このため、進化の過程において、上記のようなN人版ゲームの特徴が大きく反映される可能性がある。

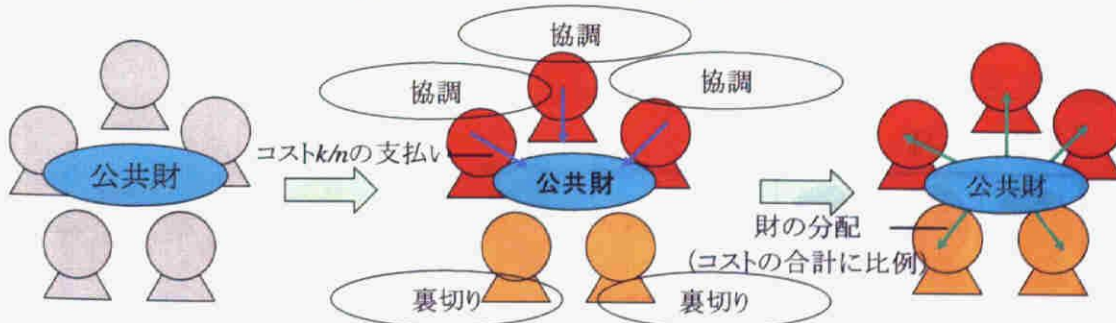


図 3.3 : N 人版ゲームの具体的状況設定の例

3.2.2 戦略の遺伝子表記

各個体は遺伝的に決定された戦略に従い、前節で定義した繰り返しゲームを行う。本研究ではひとつ前の対戦における協調者の割合に応じて、次回の手を決定する戦略を採用する。具体的には、ゲームを行うグループの大きさを n 、ひとつ前の対戦での協調者の数を c として、2つのパラメータ s ($s \in [0, 1]$ の実数)、 h (“0” または “1” の2進数) を用いて以下のとおりに定義する。

$c/n \geq s$ ならば 協調,
それ以外 裏切り.

ただし、繰り返し対戦の初回では、各個体がそれぞれ h (0 ならば裏切り, 1 ならば協調) を出すものとした擬似的な対戦履歴から協調者の割合を算出し、これを用いて初回の手を決定する。このようにして定義される戦略は、N 人版ゲームにおけるしっぺ返し戦略[Joshi 87]とみなすことができ、 s が小さいほど、また、 h が大きいほど協調的な戦略であると言える。図 3.4 は戦略の一例である。 $s=0.2$ の個体は、ひとつ前の対戦において裏切りを出していたが、その対戦での協調者の割合が $0.6 > 0.2$ であるために、現在の対戦では協調を出していることがわかる。

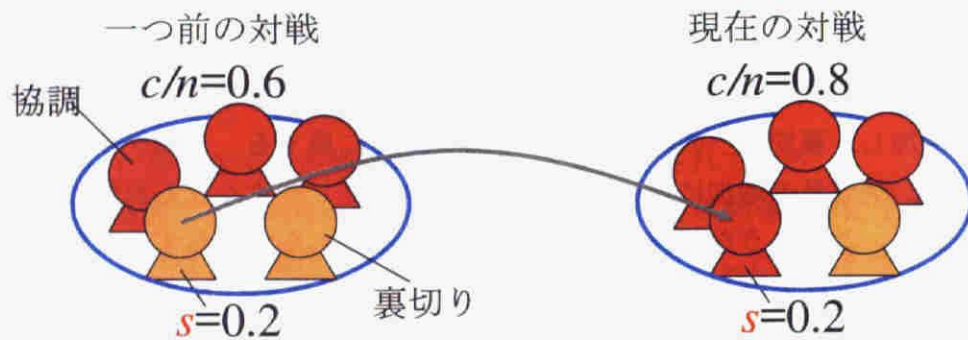


図 3.4 : N 人版ゲームでの戦略の例

3.2.3 局所的な相互作用

相互作用の規模に基づくグループ形成について解説する。集団内の各個体は、図 3.5 に示すような大きさ P の 1 次元トラス状セルに 1 個体ずつ配置される。ここで、局所性に関する一つ目の要素を導入する。各個体は、前節での戦略に関する遺伝子に加え、相互作用の規模 $i \in [2, (P-1)/2]$ を整数の遺伝子として持つ。 i は各個体が「どこまで自身の近くの個体を近傍と見なし、ゲームを行いたいのか」を表す遺伝的パラメータである。これに基づいて、各個体は図 3.5 に示すように、セル上での自身からの距離が i 以内である個体を近傍の個体とみなし、近傍の個体群で構成されるグループで N 人版ゲームを行うことを試みる。しかし、ある個体が、事前の了解なしに他の個体を一方的に自身のグループに組み入れるというのは現実社会と比較してあまり妥当な設定ではないと考えられる。そこで、少なくともグループを構成しようとする個体と他のすべての参加者との間に相互の合意が得られているという条件を付け加える。

結局、「セル x 上に存在する個体が作る相互作用に関するグループ GI_x 」を次の式で定義する。

$$GI_x = \{y \mid dist(x, y) \leq i_x \wedge dist(x, y) \leq i_y\} \quad (3.5)$$

ここで、 i_x はセル x 上の個体の相互作用の規模、 $dist(x, y)$ はセル x と y の間の距離とする。この式は、 GI_x が「 x 上の個体の近傍の個体であり、かつ、その個体自身も x 上の個体を近傍と見なしている集団」であることを表している。

例えば、図 3.6 において、セル x 上の個体が作る相互作用に関するグループは、黄色で示された領域の、 $x-3$ 上の個体を除いた $x-4, x-2 \sim x+2$ で構成される個体群となる。 $x-3$ 上の個体が含まれないのは、 x 上の個体は $x-3$ 上の個体を近傍と見なしているが、 $x-3$ 上の個体の相互作用の規模は 2 であるために、 x 上の個体を近傍と見なしていないことによる。したがって、本モデルは、ゲームを行うグループの構成員が必ずしもセル上に連続して並ばないという性質を持つ。一つの世代では、このようにして各個体によって作られる可能な N グループにすべてについて独立に、 g 回の繰り返

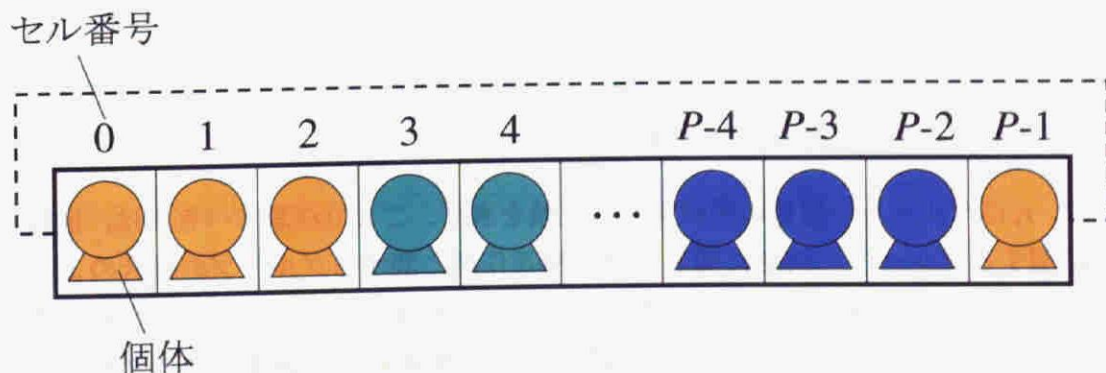


図 3.5 : 集団の構成

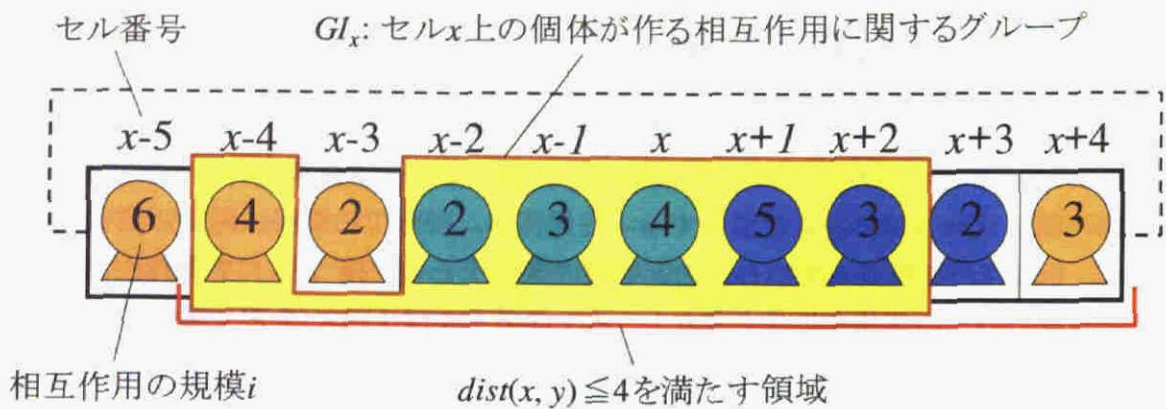


図 3.6: 相互作用に関するグループ

し対戦を行う。このとき、確率 p_n で出そうと思っていた手が反転して場に出てしまうノイズ有りの対戦を採用する。

Seoらは、N人版繰り返し囚人のジレンマゲームを行う戦略個体を1次元セル上に配置し、対戦相手の選択の際に本研究とは異なるアプローチで空間的局所性を導入した進化実験を行っている[Seo 99]。彼らは環境パラメータとして近傍のサイズを定義し、ある固定された構成員数でのゲームにおいて、そのメンバを局所的に選択する際の範囲を表すものとした。進化実験の結果、構成員数が大きい条件においても、近傍のサイズが小さくなるに従い協調行動の創発が容易になるという興味深い知見が得られている。本研究で採用する相互作用の規模は、ゲームを行うグループの構成員数自体に直接影響するという点で Seoらのモデルとは異なり、彼らとは異なる興味深い知見が得られる可能性があると言える。

3.2.4 局所的な世代交代

最後に、局所的な世代交代のプロセスについて解説する。ここで、局所性に関する二つ目の要素を導入する。まず、各個体はこれまでの遺伝的パラメータに加えて、世代交代の規模 $r \in [2, (P-1)/2]$ を整数の遺伝子として持つ。 r は各個体が「どこまで近くのセルにおける生存競争に加わりたいか」を表す遺伝的パラメータである。次世代におけるセル x 上の個体を選択するための、世代交代に関するグループ GR_x を次のように定義する(図 3.7)。

$$GR_x = \{y \mid dist(x, y) \leq r_y\} \quad (3.6)$$

ここで、 r_y はセル y 上の個体の世代交代の規模を表す。ここで注意すべき点は、世代交代に関するグループの決定に関しては、相互作用に関するグループの場合のように、相互の合意を条件としていないことである。これは、本モデルが単性生殖に基づいているために、世代交代に関しては基本的には単純な生存競争であるためであ

る。したがって、他の個体の世代交代の規模にかかわらず、各個体は自身の世代交代の規模の距離以内のすべてのセルにおける選択候補として含まれることができる。

しかし、現実世界においては、生存競争に多く身を置くことは自身の危険を増大させるという意味でコストがかかると考えられる。そこで、適応度計算では、世代交代の規模 r が大きく競争に参加する回数が増えるほど、大きなコストがかかる設定を採用する。具体的にはセル x 上の個体の適応度 F_x を次のように定める。

$$F_x = s_x \cdot (2r_x + 1)^{-\beta} \quad (3.7)$$

ここで、 s_x はセル x 上の個体のすべての繰り返し対戦における得点の平均、 $\beta \in [0.0, 1.0]$ は選択に加わることに対するコストの大きさを決める環境パラメータである。以上のようにして決定される、各セルにおける世代交代に関するグループと各個体の適応度を用いて、次世代の個体を適応度に比例したルーレット選択によって確率的に決定する。このとき、各個体の持つ遺伝子 s, h, i, r について、それぞれ確率 p_m で、定義域においてランダムに値を割り当てなおす突然変異を導入する。以上の操作を繰り返し、集団を進化させる。

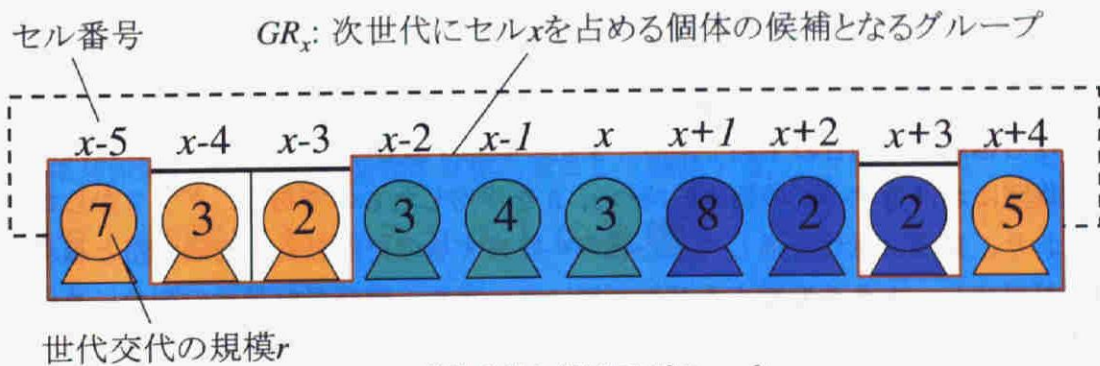


図 3.7 : 世代交代に関するグループ

3.3 空間的局所性の基本的な影響

以上のモデルを用いて、2つの空間的局所性が協調行動の創発に与える影響について議論する。はじめに、2つの局所性が進化ダイナミクスに与える基本的な影響を明らかにするため、それぞれの値を固定した実験を行った。

3.3.1 局所性の進化なしの場合での進化ダイナミクス

まず、相互作用の規模 i と世代交代の規模 r をそれぞれ、世代を通してすべての個体について同一の値 I, R を用いて実験を行った。この設定では、すべての相互作用に関するグループは、それを作る個体の近傍の $2I+1$ 個体によって構成され、すべての個体が $2I+1$ 個のグループに参加し繰り返し対戦を行う。また、世代交代は各セルについて、その近傍 $2R+1$ 個体から繰り返し対戦での平均得点に比例したルーレット選択を行った場合と同等となる。パラメータは、 $P=101, k=2.0, g=100, p_m=0.05, p_n=0.1, \beta=0.0$ 、世代数 2000 を用い、初期集団の遺伝子は定義域でランダムに決定するものとした。以後、特に断りのない限り上記のパラメータを標準のものとして用いる。

図 3.8 は、 $I=8, R=8$ を用いた場合の、協調者の割合、戦略 s, h の平均の推移を 5 世代ごとに示したものである。協調者の割合とは、各世代でのすべての繰り返し対戦における協調者の割合の平均である。また、戦略 s 及び h の平均とは、各世代において集団におけるすべての個体の持つ s, h を平均したものである。グラフからわかるとおり、集団は世代を通してある戦略へ収束するのではなく、協調戦略と裏切り戦略が混在し、大きく変動する状態が続いた。また、戦略が協調者の割合に及ぼす全体の傾向としては、戦略 s の平均が低いほど、また戦略 h の平均が大きいほど協調者の割合が高いと言える。

図 3.9 は、異なる相互作用の規模 i と世代交代の規模 r を用いて実験を行ったとき

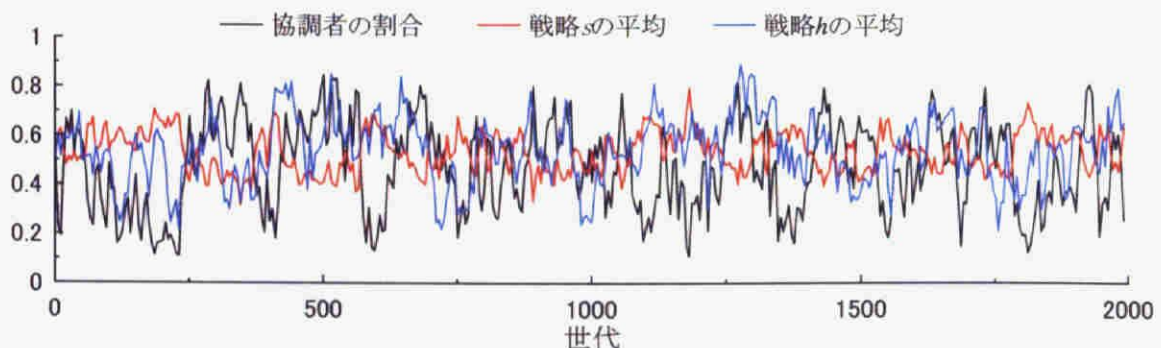


図 3.8 : $I=8, R=8$ での進化の過程

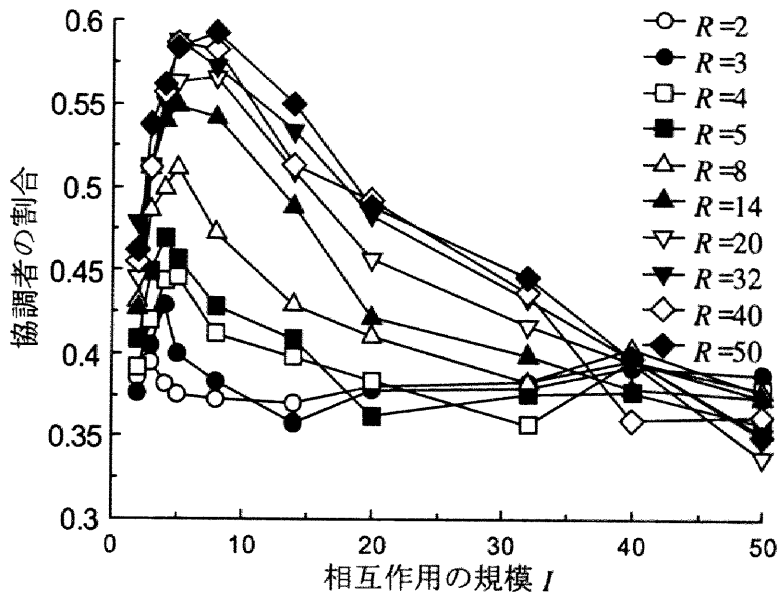


図 3.9：協調者の割合に対する I, R の影響

の世代を通したすべてのゲームにおける協調者の割合の平均を示したものである。なお、各値は 5 試行の平均である。これより、相互作用の規模の増大に伴って、協調者の割合は上昇し、あるところでピークに達し、それ以降は次第に低下していることがわかる。また、世代交代の規模の増大に伴ってピークは高くなる傾向があることがわかる。以後、このような結果をもたらす要因について、協調にかかるコスト、相互作用の規模、世代交代の規模に着目して議論する。

3.3.2 協調にかかるコストの影響

まず、図 3.9 において協調者の割合のピークをもたらす要因について考察する。このようなピークが現れる理由は、相互作用の規模 I の増加に伴う協調者の割合の増加に関する正の効果と負の効果のトレードオフによってもたらされていると言える。

正の効果をもたらす要因は、本モデルで採用する利得式の特徴であると考えられる。3.2.1 節で示したとおり、本モデルで採用した N 人版ゲームの利得式では、ゲームを行うグループのサイズが大きくなるほど、グループ内において協調にかかるコスト k が小さくなる性質がある。本節の実験では、相互作用の規模 I のとき、各グループの大きさは $2I+1$ となるため、 I が大きくなるほど、協調者の割合が増加しやすい傾向があると考えられる。

図 3.10 はコスト k の値のみ 1.5, 2.0, 2.5 に変更して、図 3.9 と同様の実験を行ったときの I の増加に伴う協調者の割合の推移を示したものである ($R=8.0$)。実際、 k が大きくなるほど協調者の割合のピークの値は低くなっていることがわかる。同時に、ピークを迎える時の I の値は k の増加と共に増加していることもわかる。これは、 k が大きい場合には、グループのサイズが大きいことで実際のコストが小さくなることが協調行動

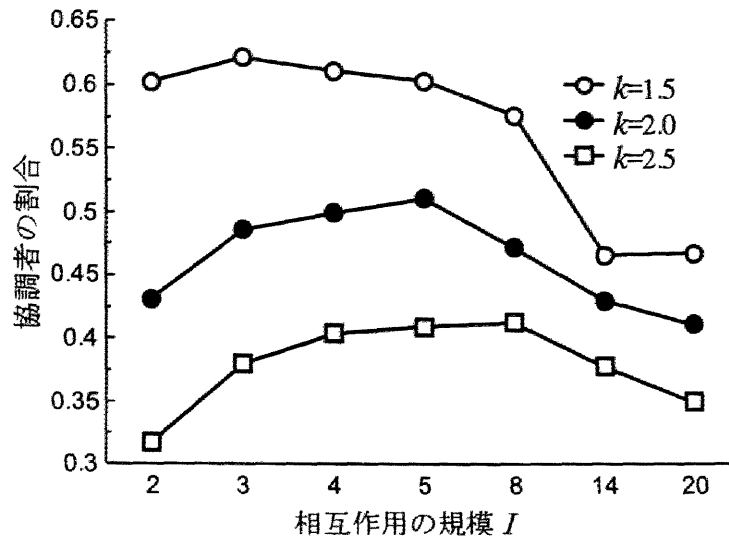


図 3.10：協調者の割合の増加に対するコスト k の影響

の創発にとって有利であったことを示している。

3.3.3 相互作用の規模の影響

一方、相互作用の規模の増加に伴う協調者の割合の増加に関する負の効果は、近傍のグループ間において生じる協調者の割合の多様性が減少することによるものであると考えられる。図 3.11, 3.12 はそれぞれ、 $I=8, R=2, 32$ の時の、ある試行での初期 200 世代における、各セル上の個体を作る相互作用に関するグループによる繰り返し対戦での協調者の割合を、濃淡（黒: 0.0, 白: 1.0）によって 5 世代ごとに示したものである。図 3.11 が示すとおり、相互作用の規模 I が小さい場合には、隣接するグループ間の協調者の割合が大きく異なっているという点で多様である。各個体の得る適応度は協調者の割合に大きく依存するため、この多様性が大きい場合には、近傍の個体間の適応度差は増加する傾向がある。

これに対し、図 3.12 より、 I が大きい場合には隣接するグループ間の協調者の割合の差が小さいことがわかる。その理由は、相互作用の規模が大きい、すなわち、ゲームを行うグループのサイズが大きいと、隣接するグループ間において、それぞれの構成員として同じプレイヤーが多く含まれるようになり、対戦結果が類似してくるためである（最も極端な例を挙げれば、 I が最も大きな値である $(P-1)/2$ のとき、すべてのグループは集団中のすべての個体をグループの構成員とするため、どのグループの対戦結果も、ノイズと初回に参照する擬似的な対戦履歴の影響を除けば同様となる）。したがって、 I が大きすぎる場合には、近傍の個体間において各個体が複数のグループに参加して対戦を行うことによって生じる選択圧がなくなってしまうことになり、これが協調者の割合の増加に対して負の影響を及ぼしていると考えら

れる. 以上のような, 相互作用の規模 I の増加に伴う協調者の割合の増加に対する正および負の効果が, 図 3.9 で示すピークを生み出していると考えられる.

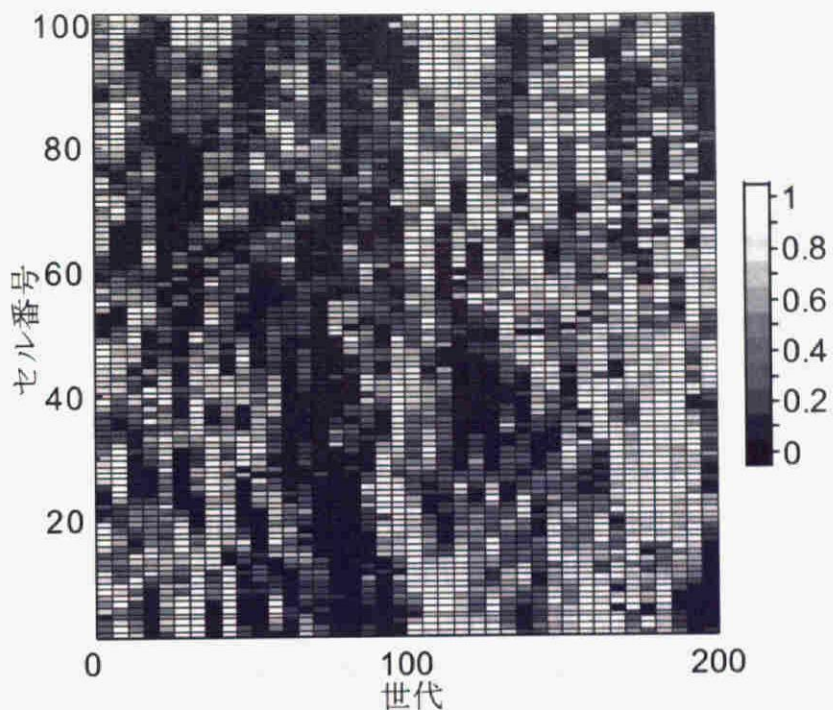


図 3.11 : $I=2, R=8$ の場合での協調者の割合の分布

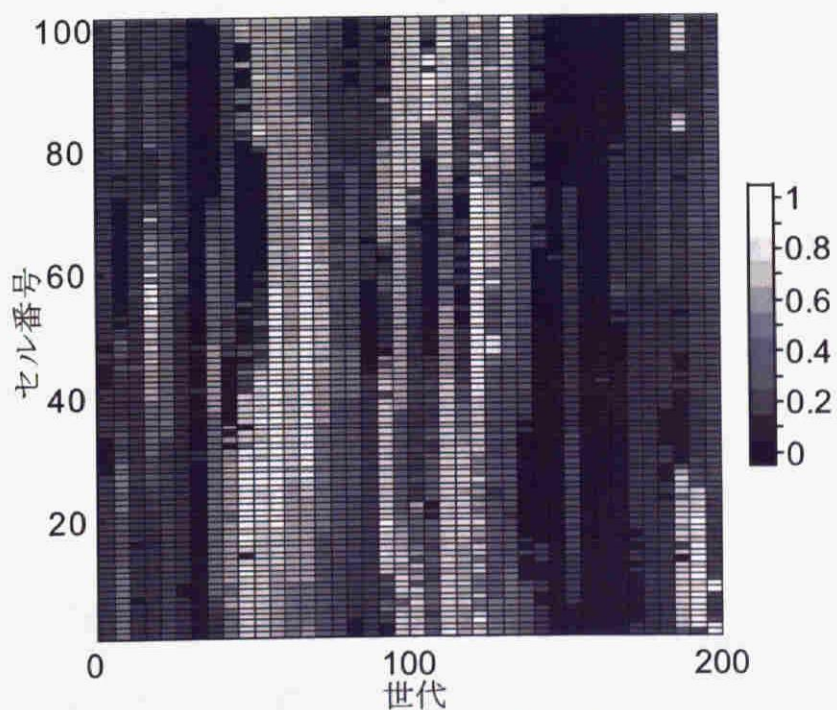


図 3.12 : $I=32, R=8$ の場合で協調者の割合の分布

3.3.4 世代交代の規模の影響

次に、世代交代の規模が集団の進化に与える影響について考察する。図 3.9 より、世代交代の規模が大きくなるに従って全体的に協調者の割合が高くなることが分かる。この要因として、世代交代の規模が集団における戦略の多様性に与える影響に注目する。

図 3.13, 3.14 はそれぞれ、 $I=8, R=2, 32$ の時の、ある試行での初期 200 世代における各セル上の個体を作る相互作用に関するグループでの協調者の割合の分布を示したものである。 $R=2$ の場合には、集団は裏切りのなグループと協調的なグループのクラスタによって分割され、かつ、世代が進むに従い頻繁に入れ替わっている。これは、局所的な世代交代が、ある戦略が集団全体を占有するのを難しくしていることを示している。それに対し、図 3.14 から、世代交代の規模の大きい $R=32$ の場合には、最適な（適応度の高い）戦略が、集団を広く急速に占有していることがわかる。このような全域的な世代交代による戦略の急速な伝播が、集団全体で適応度を増加させる方向に働くために、相互作用の規模 I に関わらず、世代交代の規模 R の増加と共に協調者の割合が増加すると考えられる。

しかし、これらの結果は、3.1.1 節で述べた、局所性が高いことは協調行動の創発を促進するという一般的な知見とは異なっている。この知見は、2 人版のジレンマゲームでは、協調的な戦略は互いに協調し合って安定なクラスタを作ることができ、それが次第に大きくなることで、集団全体を占有することができるということから得られている。しかし、本モデルでは、図 3.13 のように、局所性が高い場合には協調的な戦略のクラスタは作られるが、それが次第に拡大できるほど安定してはおらず、すぐに崩れてしまうため、上記のようなシナリオが当てはまらないと考えられる。

クラスタが容易に崩れてしまう理由の一つとしては、2 人版ゲームを採用した多くの進化ダイナミクスと比較し、本研究で採用するグループ形成の仕組みは、クラスタ内部の個体の対戦結果がクラスタ外部の個体の戦略の影響を受けやすいことが考えられる。これらより、本モデルは、世代交代の規模が小さいことで戦略群が徐々に拡大していくよりも、世代交代の規模が大きいことで協調的な戦略が急速に拡大することの方が、その後裏切りのな戦略に占有される可能性を考慮しても、集団全体における協調者の割合の増加への寄与が大きい状況であったとすることができる。

3.3.5 2つの局所性がもたらす多様性への影響

以上を踏まえ、本節で議論してきた 2 つの局所性が集団に与える影響について、集団における多様性への影響という観点からまとめると次のようになる。まず、相互作用に関する局所性については、その値が大きくなるほど、相互作用に関するグ

ループを形成する構成員が重複する為に、隣接するグループ間での協調者の割合の多様性が減少するという傾向があった。これは、集団中の隣接する個体間の適応度の多様性も減少させる働きがあると言える。一方、世代交代の規模については、その値が大きくなるほど、適応的な個体が急速に伝播できるために、戦略自体の多様性を減少させる働きがあると言える。つまり、両者の増加とも異なるレベルの多様性を減少させる働きがあるということである。

しかし、両者の協調行動の創発への影響は大きく異なり、適応度の多様性の減少は協調者の割合の増加を抑制するのに対し、戦略の多様性の減少は協調者の割合の増加を促進することが明らかになった。

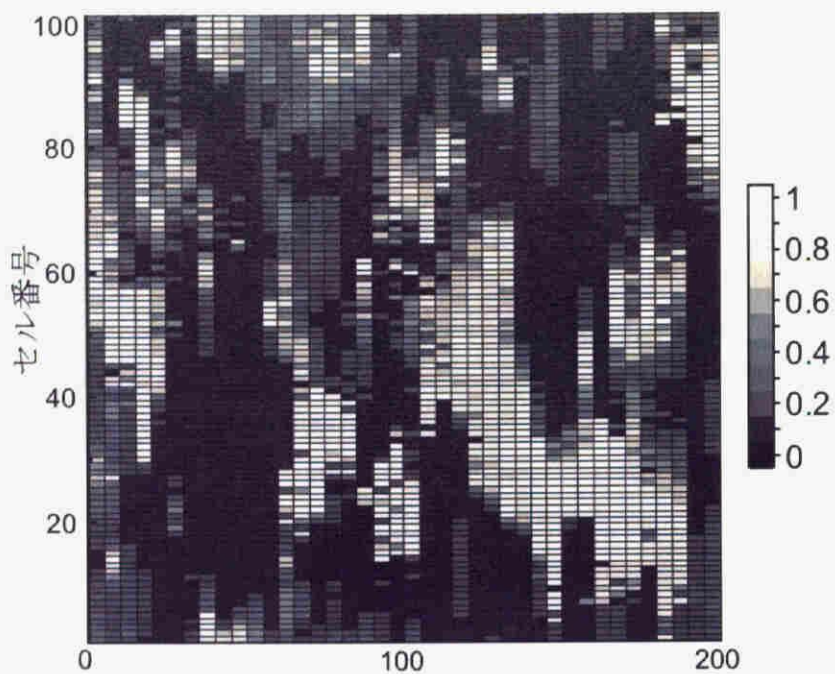


図 3.13 : $I=8, R=2$ の場合での協調者の割合の分布

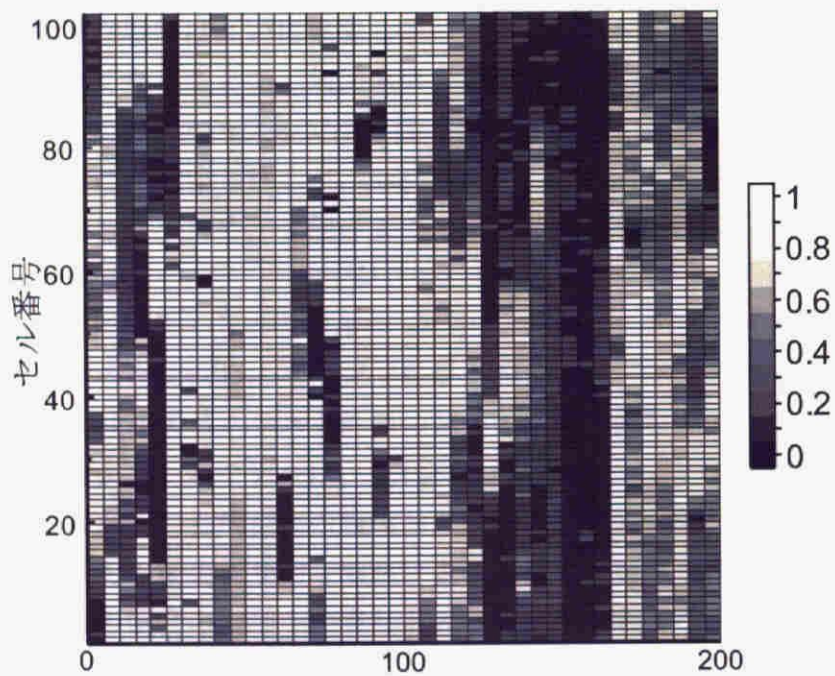


図 3.14 : $I=8, R=32$ の場合での協調者の割合の分布

3.4 空間的局所性の進化

前節では、相互作用の規模と世代交代の規模が、協調行動の創発に関して大きく異なる影響を与えることを明らかにした。本節では、これらの局所性を遺伝的パラメータとして進化させた場合における集団の挙動について明らかにする。本節では特に、2つの局所性それぞれが進化することによる影響をより鮮明にするために、片方ずつ進化を導入して実験を行う。

3.4.1 相互作用の規模の進化

はじめに、相互作用の規模の進化を導入して実験を行った。具体的には、3.2.3 節で示したとおりに、各個体が遺伝的に保持する相互作用の規模 i に従って相互作用に関するグループの形成を行うものとした。これにより、集団内には異なるサイズのゲームを行うグループが同時に存在することになる。図 3.15 は $R=8$ を用いた場合の協調者の割合と相互作用の規模の平均、および、グループの大きさの平均を示したものである。相互作用の規模の平均、グループの大きさの平均とは、各世代において集団におけるすべての個体の持つ相互作用の規模 i 、すべての相互作用に関するグループの大きさを平均したものである。この場合においても、相互作用の規模の進化なしの場合と同様、協調戦略と裏切り戦略が混在し、協調者の割合が大きく変動する状態が続いた。ここで注目すべき点は、相互作用の規模が、前節において協調者の割合のピークを示した値に収束するのではなく、世代を通して大きく振動を繰り返したことである。

図 3.16 は、同じ条件でのある試行における、世代ごとの協調者の割合とグループの大きさ、相互作用の規模の平均の相関を 5 世代ごとに示したものである（図中の直線は最小二乗近似による近似直線）。同図は、これらに緩やかな相関（相関係数は 0.48）があることを示している。図 3.15 からわかるように、その進化ダイナミクスを予測することは容易ではないものの、これは、協調的な戦略にとって、グループの大きさを広げることで協調にかかるコストを減少させることができることに

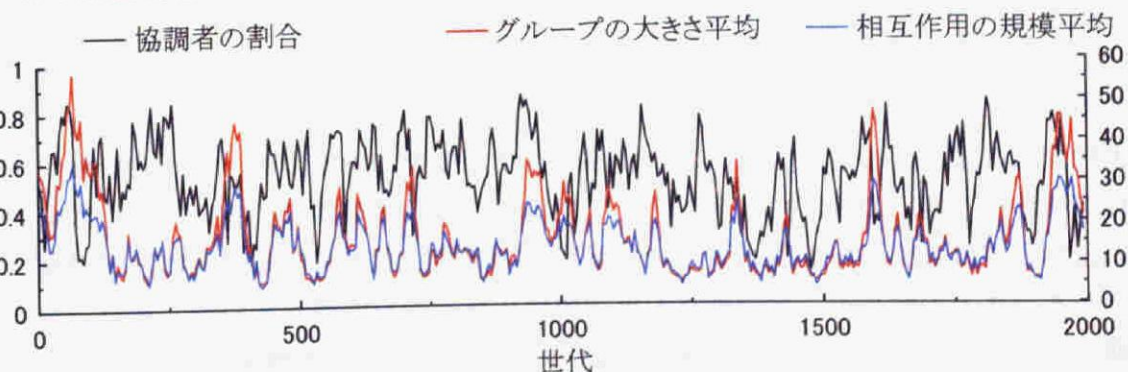


図 3.15：相互作用の規模 i を進化、 $R=8$ の場合での進化の過程

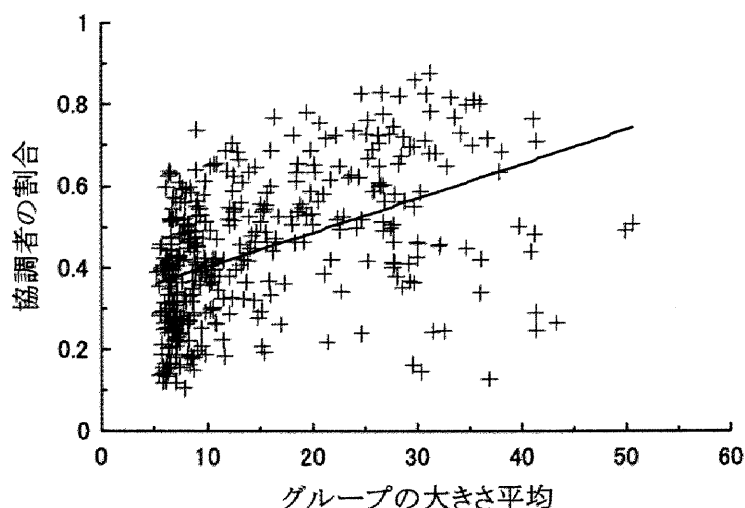


図 3.16：協調者の割合とグループの大きさの相関

起因すると考えられる。つまり、もし、協調的な戦略が集団中に局所的にでも広まることができた場合には、相互作用の規模は大きくなる方向へ進化しやすいと言える。このことは更に協調的な戦略が広まることを促進するために、協調者の割合の増加と相互作用の規模には相関が見られると考えられる。しかし、協調的で相互作用の規模の大きい個体があまりに増加しすぎると、逆に裏切りの戦略が容易に侵入できてしまう状況を作り出してしまうために、進化のシナリオは安定せず図 3.15 のような挙動を示すと考えられる。

このような相互作用の規模の複雑な進化ダイナミクスが協調行動の創発を促進するかどうかは、環境パラメータとして用いる世代交代の規模に依存する。図 3.17 は $R=2, 8, 14, 20, 26, 32, 38, 50$ の場合の世代を通したすべてのゲームにおける協調者の割合の平均を表したものである。各値は 5 試行の平均である。同時に、比較のために、相互作用の規模の進化なしでの実験において、世代交代の規模 R を用いて I の値を変えて実験を行ったときの、協調者の割合が最も高かったときの値を示している。つまり、図 3.9 において、各 R の条件におけるピークの値を比較として示しているということである。図 3.17 において注目すべきは、世代交代の規模 R が 14 以下のとき、相互作用の規模が進化する場合の方が進化しない場合よりも協調者の割合が高いという点である。逆に、 R が 20 以上の条件では、相互作用の規模の進化は協調行動の創発に寄与していないこともわかる。

世代交代の規模が小さい場合において協調行動が促進されるのは、局所的な世代交代が行われる集団では、3.3.4 節で示したとおり、ある戦略が集団中に大きく広まるのは容易ではないが、前述のような相互作用の規模が大きくなることによる協調にかかるコストの減少が、協調戦略が集団中に広まるのを促進していることが一つの理由として挙げられる。表 3.2 は各世代交代の規模 R を用いて実験を行ったときの、相互作用の規模 i の平均と、進化なしの実験で協調者の割合が最大となったときの相

相互作用の規模 I をグループの大きさと共に示したものである。これより、実際、進化なしの場合において協調者の割合が最大という意味で最適なグループの大きさと比較して、相互作用の規模の進化は約 2 倍のグループの大きさへと押し広げており、コストが比較的小さい条件で集団が進化していることがわかる。

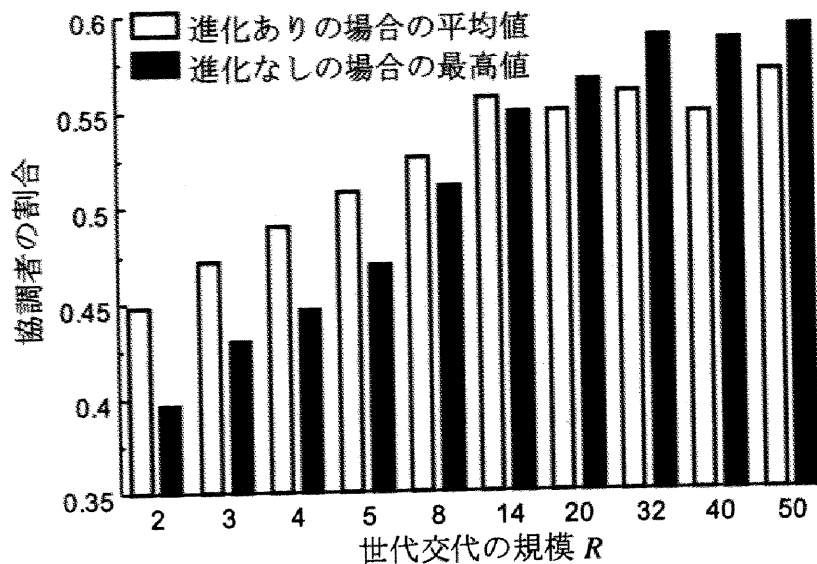


図 3.17：相互作用の規模の進化を導入したときの協調者の割合

表 3.2：相互作用の規模の進化を導入した場合の平均値

世代交代の規模 R	相互作用の規模 i の平均 (グループの大きさ平均)	進化なしの実験で協調者の割合が最大となった時の相互作用の規模 I (グループの大きさ)
2	16.9 (19.6)	40.0 (81.0)
3	16.4 (19.4)	4.0 (9.0)
4	16.3 (19.2)	5.0 (11.0)
5	15.9 (18.7)	4.0 (9.0)
8	15.6 (18.5)	5.0 (11.0)
14	16.9 (20.8)	5.0 (11.0)
20	17.9 (23.0)	8.0 (17.0)
32	18.4 (23.7)	5.0 (11.0)
40	17.6 (22.4)	5.0 (11.0)
50	18.6 (24.3)	8.0 (17.0)

3.4.2 世代交代の規模の進化

次に、世代交代の規模の進化を導入して実験を行った。具体的には、3.2.4 節で示したとおりに、各個体は遺伝的に保持する世代交代の規模 r に基づいて各セルにおける世代交代を行うためのグループに参加するものとした。図 3.18 は世代交代の規模 r を進化させ、 $I=8$ 、 $\beta=0.4$ としたときの世代交代の規模と協調者の割合の推移を示したものである。同図より、相互作用の規模の進化の場合とは異なり、世代交代の規模は、多少の増減はあるものの、ある一定の値に収束する傾向があることがわかる。

図 3.19 は $I=8$ に固定し、 β を 0.0~1.0 まで 0.1 刻みに設定して実験を行ったときの、世代を通しての世代交代の規模 r の平均とすべての対戦での協調者の割合の平均を示したものである。なお、各値は 5 試行の平均である。同図より、世代交代の規模が収束する値は β に大きく依存することがわかる。まず、 $\beta=0.0$ から値が増加するに従い、協調者の割合はばらつきながらもゆるやかに減少する傾向があった。世代交代の規模平均については、 $\beta=0.0$ のときは約 45 と極めて大きな値へと収束した。これは、この条件では世代交代の規模が大きいことによるコストが全くかからないために、多くの選択に参加したほうが明らかに有利であることによる。

更に、 β が増加するに従って、約 0.3 までは緩やかに減少した後、0.7 までは大きく減少し、それ以降は再び緩やかに減少することがわかる。これは、コストが大きすぎる場合には、多くの選択に参加するよりも近傍の少ないセルにおける選択のみに参加した方が有利であることを示している。また、 $\beta=0.3\sim 0.7$ の場合にはこのメリットとコスト β とのつりあいが取れる条件が、世代交代の規模が中程度の値であったことを示している。

以上より、世代交代の規模は環境条件によって決定されるコスト β に依存して、進化の過程で一定の値に収束することが明らかになった。

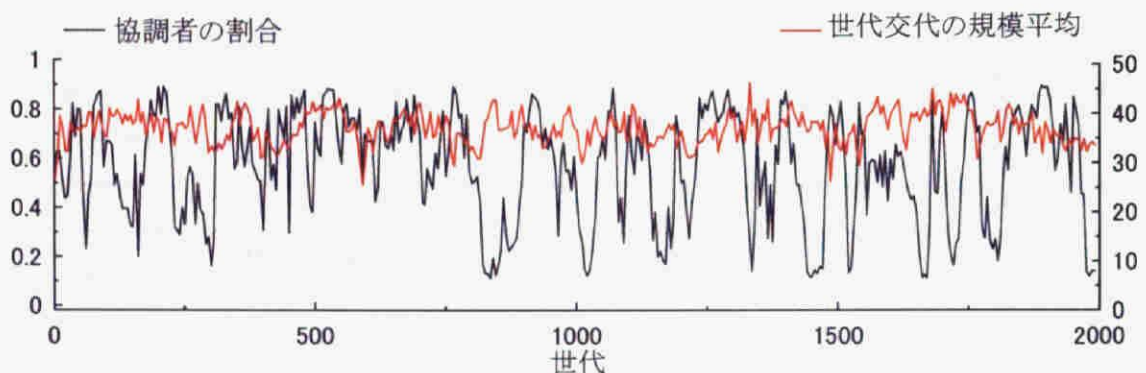


図 3.18：世代交代の規模 r を進化、 $I=8$ 、 $\beta=0.4$ の場合での進化の過程

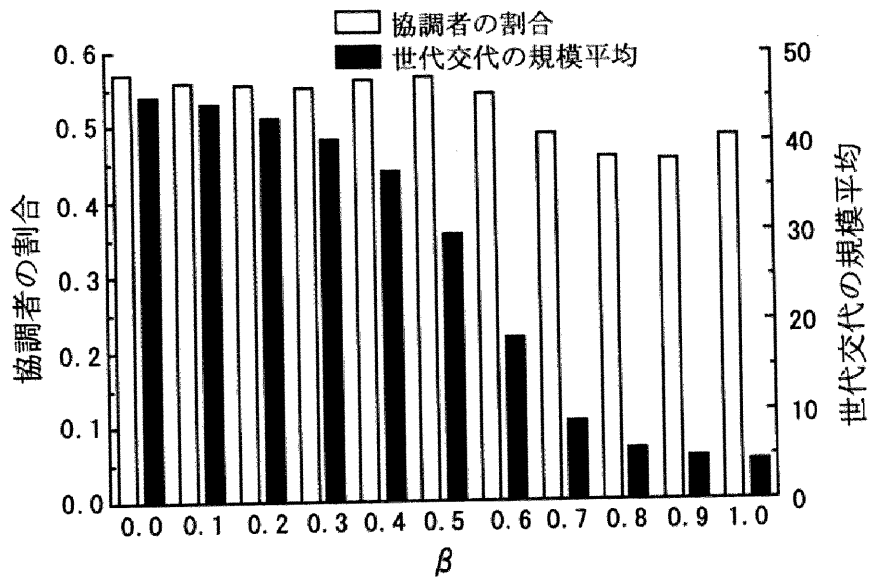


図 3.19 : 協調者の割合と世代交代の規模の進化に対する β の影響

3.5 まとめ

本章では、利己的集団における協調行動の創発に関して、空間的局所性の与える影響について知見を得るために、空間的局所性とその進化を導入した N 人版繰り返し囚人のジレンマゲームにおける戦略の進化モデルを構築し、進化実験を行った。このとき、N 人版繰り返し囚人のジレンマゲームの戦略を 1 次元セル上に配置し、各個体がゲームの戦略の他にどれくらい近傍の個体と相互作用・世代交代を行うかという相互作用の規模・世代交代の規模を個別に設定可能なモデルを採用した。その結果、相互作用の規模の増大に伴って協調関係の創発にピークが存在することが明らかになった。これは、本モデルで採用した N 人ゲームの利得式の特性である、構成員数が多くなるほど協調することにかかるコストが小さくなることと、相互作用の規模が大きくなることによるグループ間の協調者の割合の多様性の減少とのトレードオフによるものであった。また、世代交代の規模は大きいほど協調集団への進化に有利であることが判明した。これは、全域的な世代交代は、協調的な戦略がすばやく集団全体に広まるのを容易にするためであると考えられる。これらの結果は、繰り返し囚人のジレンマゲームの戦略進化において一般的に言われる、局所性が高いほど協調行動の創発が促進されるという知見とは異なるものであり、興味深いと言える。

次に、相互作用の規模、世代交代の規模を各個体が戦略の一部として遺伝子に保持し、進化する設定を用いて実験を行った。はじめに、相互作用の規模の進化を導入して実験を行った。その結果、相互作用の規模は、ある一定の値に収束することではなく、世代を通して増加減少を繰り返した。また、このような進化が協調行動の促進に寄与するかどうかは環境パラメータとして与えられる世代交代の規模に依存し、これが小さい場合には、相互作用の規模の進化なしの場合の最適値よりも、進化ありのほうが協調者の割合が増加することが明らかになった。一方、世代交代の規模の進化を導入した場合には、世代交代の規模はある値に急速に収束することが判明した。収束する値は、選択に参加すること自体にかかるコストを表す環境パラメータに依存し、コストが増加するに従って収束する値は減少していった。以上より、相互作用・世代交代の規模の進化ダイナミクスは大きく異なることが明らかになった。

以上の議論で得られた知見は、原始的・社会的集団におけるグループ形成のメカニズムやその起源の解明において重要であるだけでなく、近年注目されるマルチエージェント進化ロボティクスといった自律分散系への応用が期待できる[石田 99]。例えば、Embodied Evolution[Watson 99b]のような、実ロボットを用いてその行動規則を進化的に獲得することを考えた場合、ロボット間のタスク遂行を通じた局所的な相互作用はもとより、その進化自体も個体間の局所的なやり取りによって行われる

状況が想定される。世代交代の規模の進化実験で得られた知見は、このような局所性に関わるパラメータ自体も進化メカニズムに組み込むことによって、物理的制約などによって与えられるコストなどと折り合いのついた最適な値に調節が可能であることを示している。さらに、相互作用の規模の進化実験の結果は、局所性に関するパラメータが進化し、時間経過と共に変動すること自体が、系全体により適した状態をもたらす可能性があることを示唆していると言える。

また、今後の研究の展開の一つとして、本モデルにおける相互作用・世代交代の規模の進化をネットワーク構造の進化と見なした議論が挙げられる。近年、現実世界に存在するネットワーク構造の特徴や、その構造の自己組織的な進化・発展に関する研究が注目されている。本研究で議論してきた相互作用・世代交代に関するグループの形成は、集団内でネットワークを構築することに他ならない。このような視点から相互作用・世代交代の規模の進化ダイナミクスに関して議論することは、新たな知見をもたらす可能性がある。

例えば、Watts らが提案した、現実世界に存在するネットワークの特徴の一つであるスモールワールド構造に関する指標（ネットワーク中のノードの組についての最短パス長の平均と、近傍の派閥度）などを用いた特徴量を、相互作用や世代交代の規模が作るネットワークに適用することなどが挙げられる[Watts 98]。彼らによれば、ネットワーク上に置かれたしっぺ返し戦略を行う個体とその近傍の個体と N 人版囚人のジレンマゲームを繰り返し行う設定では、近傍の派閥度が小さいほど、また、ショートカット（最短パス長を大きく減少させるようなリンク）の数が増加するほど、裏切りの集団へと収束しやすい環境となることを指摘している（戦略としてパブプロ戦略を用いた場合にはショートカットが増加するほど若干協調戦略が進化しやすいということを同時に指摘している点は興味深い）。また、全面裏切り戦略としっぺ返し戦略が混在して進化する場合には、ショートカットが多いことが必ずしも協調集団への進化を妨げないことも指摘している[Watts 99]。

しかし、これまでに行った基礎的実験では、本モデルは、ノード数に対してリンク数が多い、近傍のノード間の結合が常に強い、相互作用の規模が増加するとリンク数も大きく増加するといった性質のために、上記のような特徴量と協調行動の創発に関して明確な知見は得られていない。本研究の成果を踏まえ、ネットワーク構造の進化に対してより自由度を与えたり、逆にリンク数に制約を加えたりするなどの変更をモデルに施すことによって、局所性の進化に関して新たな知見が得られることが期待できる。

第4章 行動多様性に対する情報共有の影響とその適応性

これまでの2つの研究事例では、ゲーム論的状況設定を用いた抽象的なモデルでの議論であった。本章では、これまでと比較して具体的な状況設定を用いて、そこに生じる創発的ダイナミクスについて論ずる。

大規模テーマパークなどのイベント会場において、携帯情報端末などを用いて来場者に混雑情報を提供し、混雑を緩和する試みが検討されている。しかし、例えば、ある時点で「あるアトラクションが空いている」という情報を過剰な数の来場者が共有し、それに従って行動した場合、結果としてそのアトラクションが混雑してしまうことがありうる。これは、集団においてある情報を共有することが行動の画一化を招くような場合、必ずしも系全体にとって良い結果をもたらすとは限らないことを示す一例である。このような、系に生じる多様性の適応性やその起源・維持のメカニズムの解明は、複雑系科学における主要な研究テーマの一つである。

本章では、このような群集行動に対する情報共有の影響、特に、個体間で共有される情報の多様性と行動の多様性の間に生じる相互作用に注目し、両者の関係がもたらす創発的ダイナミクスを明らかにすることを目的とする。このために、イベント会場における混雑解消のための混雑情報提供を、集団における行動の多様性に対する情報共有の影響とその適応性に関する具体的な状況設定として取り上げる。来場者が、混雑情報と各自の嗜好に基づいて複数のアトラクションを観賞してまわる抽象モデルを用いてマルチエージェント・シミュレーションを行い、情報共有が与える群集行動への影響について議論する。

はじめに、会場内の施設の配置が混雑の偏りに与える基本的な影響を明らかにする。次に、来場者に対して混雑情報の提供を行い、来場者がそれに従って行動する状況を設定した場合、集団全体の混雑にどのような影響を及ぼすかについて解析を行い、その効果を明らかにする。更に、来場者数を増やして長い期間シミュレーションを行うことで、時間経過と共に混雑共有の効果がどのような変化を見せ、系全体にどのような挙動をもたらすかについて議論する。本章では、特に、あるアトラクションにおける待ち行列の突発的な発生という動的な現象に注目し、情報共有が行動多様性にもたらす動的な影響について詳細に議論する。

本章の構成は次のとおりである。4.1節では、研究の背景と目的について述べる。4.2節では、構築したモデルについて解説する。4.3節では、施設配置に起因する混雑の偏りに対する情報共有の影響について明らかにする。4.4節では、情報共有の効果の時間変化が系全体にもたらす影響について議論する。最後に、4.5節で、これらの結果をまとめる。

4.1 背景と目的

4.1.1 情報の多様性と行動の多様性

集団において、ある情報を共有することが、その情報の効果や価値を低下させてしまう場合がある。例えば、カーナビゲーションシステムを利用した道路の渋滞情報の提供が渋滞緩和のために行われているが、もし、「ある一般道が渋滞している」という情報を過剰な数の運転者が共有し、渋滞を避けようとして裏道を選択した場合、結果としてその裏道が混雑してしまうという状況が考えられる。株取引や流行などに関する情報においても上記のような現象は容易に見出しうるものであり、情報というもののひとつの本質であると考えられる。

このような現象が発生する要因は、集団において共有される情報が、ある限られた資源への集中を促すものであることだと考えられる。この場合、ある行動を行う（ある資源を選択する）ことのメリットが、集団内で同様にそれを行う個体の頻度に対し、負の依存関係にあると見なすことができる。現実世界においては、この対象となる資源は、物質的資源のみならず、時間や空間、生態学的地位など様々な形で存在する。このような状況では、いかにして集団内の行動の多様性を適度に維持するかが全体の適応性が維持されるための鍵であり、前述の例は、情報共有によって行動の多様性が過度に減少してしまうケースであると考えられる。系に生じる多様性の適応性やその起源・維持のメカニズムは、複雑系科学における主要な研究テーマの一つとして議論されてきた。

社会的集団における、有限の資源をめぐる複数の主体間の競合に関しては、El farol bar 問題[Arthur 94]や Minority game[Challet 97]をはじめとして様々な領域で議論されている。Minority game とは、複数のエージェントによって構成される集団において、各エージェントがステップごとに、同時に 0 または 1 の選択肢のうちどちらかを選択し、選択者数が少ない方を選択したエージェントが勝者となる繰り返しゲームである。2つの選択肢は、例えば、前述の裏道を使う・使わない、株取引における売り・買いといった、二者択一の状況において少数派を選択した方が有利になる状況が当てはまる。このような状況においては、系全体の過去の履歴から次のステップにおける系全体の状態を正しく予測する合理的な推論は存在せず、エージェントは帰納的な方法によって自身の行動規則を更新し、集団内の行動の多様性を保たなくてはならない。Savit らは、このような状況で各エージェントがランダムに選択肢を選んだ場合よりも系全体が効率的な状態が、集団のサイズと履歴の長さを調節することで実現可能であることを示した[Savit 99]。これらは、各主体が持つ行動規則の多様性が系全体の挙動に大きな影響を及ぼすことを示している。

また、有限の資源獲得をめぐる競合に関して、各主体が保持する情報の多様性に

注目した研究がなされてきた。Hoggらは、2種類の資源の獲得をめぐって、時間遅れやバイアスを伴った資源に関する情報に基づいて各主体が行動を選択する自律分散系において、その大域的な挙動がカオス的で不安定になる場合があることを指摘した上で、各主体が得る情報の時間遅れやバイアスの度合いを即時の報酬に基づく進化的手法によって調整することで、カオスを抑制する集団の構成比を獲得可能であることを示した[Hogg 91]。この手法は、ネットワークにおけるパケットのルーティングなどにおいてその適用が検討されている[山崎 98]。

また、赤石らは、資源収集問題を対象としたマルチエージェントモデルにおいて、情報の共有に伴う行動の集中化と多様化の観点から、環境からの資源情報獲得やエージェント間の資源情報の授受の際に発生する誤認の適応性について議論している[Akaishi 02]。彼らは、特定の行動を促進する情報を環境から直接得る際に生じる直接誤認や主体間の情報のやり取りにおいて生じる間接誤認が、個体が保持する情報を多様化し適応的となりうる条件が存在することや、両者の効果は相補的になりうるなどを定量的に示した。これらは、各主体が保持する情報の多様性の維持が、系全体の安定性や適応性にとって重要であることを示していると言える。

このように、集団の様々なレベルで生じる多様性の維持やその適応性について、多くの議論がなされてきた。

4.1.2 目的

では、集団中に異なるレベルの多様性が複数存在した場合、それらはどのように影響し合うのであろうか。例えば、前述の情報と行動に関する多様性は、直接的または間接的に、どのような影響を相互に及ぼすのだろうか。集団において異なる多様性の存在を想定することはより現実的な問題設定であり、その相互作用のメカニズムを探ることは議論すべき重要な問題設定である。そこで、本研究では、集団において共有される情報と行動の多様性の両者を考慮することができるより具体的な状況設定として、イベント会場における混雑情報提供による混雑緩和を取り上げ、抽象モデルによるマルチエージェント・シミュレーションを通して、両者の影響について知見を得ることを目的とする[鈴木 02b, 02c, 投稿中], [Suzuki accepted b]。

博覧会や人気テーマパークなどの大規模なイベント会場において、来場者の作る混雑は大きな問題のひとつとなっている。特に、施設を観賞するための待ち時間は、来場者の満足感に大きく影響するだけでなく、イベントを開催する側にとっても効率的な運営や安全確保のために注目すべき点である。そこで、イベント会場において、携帯情報端末などを用いたリアルタイムな情報提供による混雑緩和が検討されつつある。ここで、会場内の各施設の収容能力を限られた資源、端末による情報提供を来場者間の情報共有、各来場者の嗜好を行動規則と見なすと、前述と同様な構造を持つ具体的な状況設定として捉えることができる。

これらの設定を用いて、具体的には、次のような問いについて明らかにすることを目的とする。

- 会場内における施設（アトラクション）の配置が施設ごとの混雑（待ち行列の長さ）にどのような影響を与えるか？
- 来場者に対する混雑情報提供によってもたらされる来場者間の情報共有が、配置による混雑の偏りをどのように緩和し、行動の多様性を増加させるか？また、偏りはどのような条件で最も緩和されるか？
- 上記のような情報共有の効果は、時間経過と共に変化するか？変化するとすれば、それが集団全体の行動の多様性にもたらす影響とはどのようなものか？

このために、イベント会場に設置された複数の施設を多数の来場者が各自の嗜好に基づいて観賞してまわる抽象モデルを構築した。本モデルでは、各エージェントは混雑情報を共有する場合としない場合が確率的に決定される。また、来場者の行動はそれまでの行動履歴と嗜好に依存して決定されるために、来場者が同一の情報を受け取っても実際の行動が異なる。したがって、前述の共有する情報に関する多様性と行動に関する多様性の両方を含んだモデルであると言える。

なお、辺見は、遊園地における混雑情報と入場者の行動について、混雑情報の提供が集団の挙動に与える影響についてシンプルなマルチエージェントモデルを用いて議論しているが[辺見 02]、本研究では、各アトラクションの配置が混雑にもたらす影響を前提にして、配置によって生じる混雑の偏りが情報提供によってどのように緩和され、その効果が時間発展と共にどのように変動するかについて、情報と行動の多様性に注目して議論する。

4.2 イベント会場における混雑情報提供モデル

4.2.1 イベント会場

イベント会場における施設配置と混雑情報の提供が来場者の作る混雑にどのような影響を及ぼすかについて議論するために、次のような抽象モデルを構築した。まず、イベント会場の敷地を $W_x \times W_y$ の 2 次元のセル状平面で表現する。会場には N_a 個のアトラクションが存在する。平面左上端のセルを入り口とし、来場者（以下エージェントと呼ぶ）は入り口に配置されることにより会場に入場する。会場に一度に存在することのできるエージェント数には制限 (N_m 人) があり、制限を越えない範囲で全エージェント数 (N_p 人) に達するまで入り口にエージェントが 1 ステップにつき最大 N_e 人ずつ配置される。

4.2.2 エージェントの行動

各エージェントは、入場時に N_a カ所のアトラクション A_i ($i=0, 1, \dots, N_a-1$) それぞれについて確率 P_i でそのアトラクションを観賞するかどうか決定するものとする（観賞すると決めたアトラクションを、観賞予定のアトラクションと呼ぶ）。 P_i は各アトラクションの人気度をあらわす。また、確率的に決定された観賞予定アトラクションの組み合わせの違いは、エージェントごとの嗜好の違いと捉えることができる。各ステップでは、はじめに会場内のすべてのエージェントに対して、重複を許さずにランダムな番号を割り当て、その番号の小さいエージェントから順に、次に挙げる 3 つの“モード”を遷移しながら行動するものとする。入場時のエージェントは“移動モード”であり、観賞予定のアトラクションのうちランダムに選択したものをはじめの訪問先とする。移動モードでは、自身を中心とした 9 近傍のセルのうち、会場の外、及び、アトラクション以外で他のエージェントの移動先となっているセルの数を n とし、確率 $1-(n+1)/10$ で移動を行うかどうかを決定する（それ以外の場合にはそのステップでは移動しない）。これは、エージェント間に生じる干渉を表すものである。移動する場合には、次の式に従って x 及び y 座標を更新する。

$$x \leftarrow x + r \operatorname{sgn}(A_{ix} - x) \quad (4.1)$$

$$y \leftarrow y + r \operatorname{sgn}(A_{iy} - y) \quad (4.2)$$

ここで、 A_{ix} , A_{iy} は現在向かっているアトラクション A_i の x 及び y 座標、 $r \operatorname{sgn}(k)$ は、「確率 p_r で $-1, 0, 1$ のうちからランダムに値を返し、それ以外（確率 $1-p_r$ ）の場合は、 $k=0$ ならば 0 , $k>0$ ならば 1 , $k<0$ ならば -1 の値を返す」関数とする。これは、確率的に揺らぎつつ、各座標軸方向について目的地へ近づく方へ移動することを表してい

る（ただし、新たな座標を算出した結果、座標が会場からはみ出している場合には、はみ出した座標方向への移動は行わないものとする）。また、既に他の移動済みのエージェントが移動先のセルに存在している場合には、そのステップでは移動せず、現在位置を移動先とする（ただし、移動先がアトラクションの場合は常に移動可能とする）。このため、他の移動済みのエージェントが既に現在位置に存在しており、かつ自身が移動できない場合には、同一のセルに複数のエージェントが存在することになる。また、移動方向に依存した速度差をなくすために、上記の手順によって決定された移動先が現在位置から斜め方向の位置である場合、更に確率 $1-2^{-0.5}$ で移動を取りやめるものとする。

アトラクションに到着した（アトラクションと同一のセルに移動した）エージェントは、行列の最後尾に並んだものとして“待機モード”に入り、その場で待機する（ただし同一ステップにおいて複数のエージェントがアトラクションに到着した場合、その並び順はステップ終了時にランダムに決定する）。各アトラクションでは、最大 C_i 人のエージェントが同時に観賞可能であり、この制限を越えない範囲で並んだ順に“観賞モード”に入り、 L_i ステップの間観賞する（実際には何もしない）。本研究では、テーマパークや博覧会会場など、大規模で常に施設を観賞するための待ち行列が観察されるような状況を考えるため、アトラクションの収容能力 C_i と比較して入場者が多い設定を採用する。したがって、アトラクションを観賞するために多くの待機モードのエージェント（待機エージェント）が存在することになる。

4.2.3 混雑情報の提供と情報共有

観賞を終えたエージェントはそのアトラクションを観賞予定のアトラクションから除外し、残りの観賞予定のアトラクションの中から次に目指すものを決定して、再び移動モードに入る。ここで、来場者はそれぞれ入り口で携帯情報端末を渡されていると仮定する。端末には、来場者がアトラクションを観賞し終えるごとに確率 r で混雑情報、すなわち各アトラクションの現在の待機エージェント数が提示される。その場合、来場者はその情報を見て、残りの観賞予定のアトラクションの中からその時点で最も待機エージェント数が少ないものを次の行き先として選ぶものとする。それ以外（確率 $1-r$ ）の場合には混雑情報は提示されず、残りの観賞予定のアトラクションの中から直線距離で最も近くにあるものを選ぶものとする（この状況は、近隣のアトラクションに関する情報を受け取った場合とみなすこともできる）。なお、一度行き先を決定した後はそこに到着するまで行き先を変更しないものとする。

本モデルでは、混雑情報の提供確率 r が大きいほど、その情報を多くのエージェントが共有することになる。したがって、情報を共有する・しないという極端な例であるが、確率 r は集団内で共有する情報の多様性を象徴するものと捉えることができる。観賞予定のアトラクションがなくなったエージェントは会場を去ったもの

として消去される。シミュレーションはすべてのエージェントが会場を去った時点で終了する。

4.3 混雑に対する情報共有の影響

4.3.1 アトラクションの配置の影響

図4.1は、シミュレーションの実行画面の一例を示したものである。図中、黒い長方形で囲まれた緑色の領域が会場、番号付きの白い円に囲まれた正方形がアトラクション、小さな黄色い円が来場者エージェントを表している。画面左上端の入り口からエージェントが入場し、目的のアトラクションへと移動している様子がわかる。また、アトラクションの右側には、そのアトラクションにおける待機エージェント数が棒グラフで表されており、棒が長いほど多くの来場者がそのアトラクションを觀賞するために並んで待っていることを表している。画面右側には、現在のステップ数、エージェントのモードなどの情報が表示されている。

本研究では、アトラクションの配置と混雑情報共有のみが混雑に与える影響を明らかにするために、各アトラクションの人気度、觀賞時間、収容能力をすべて同一の値 $P_i=0.7$, $L_i=30$, $C_i=10$ に設定した上で、図4.1に示すような配置を用いて実験を行った。また、その他のパラメータ設定として $W_x=70$, $W_y=50$, $N_a=10$, $N_p=2000$, $N_e=2$, $N_m=1000$, $p_r=0.05$ を用いた（以後、特に断りのない限り上記パラメータを標準の値として用いるものとする）。

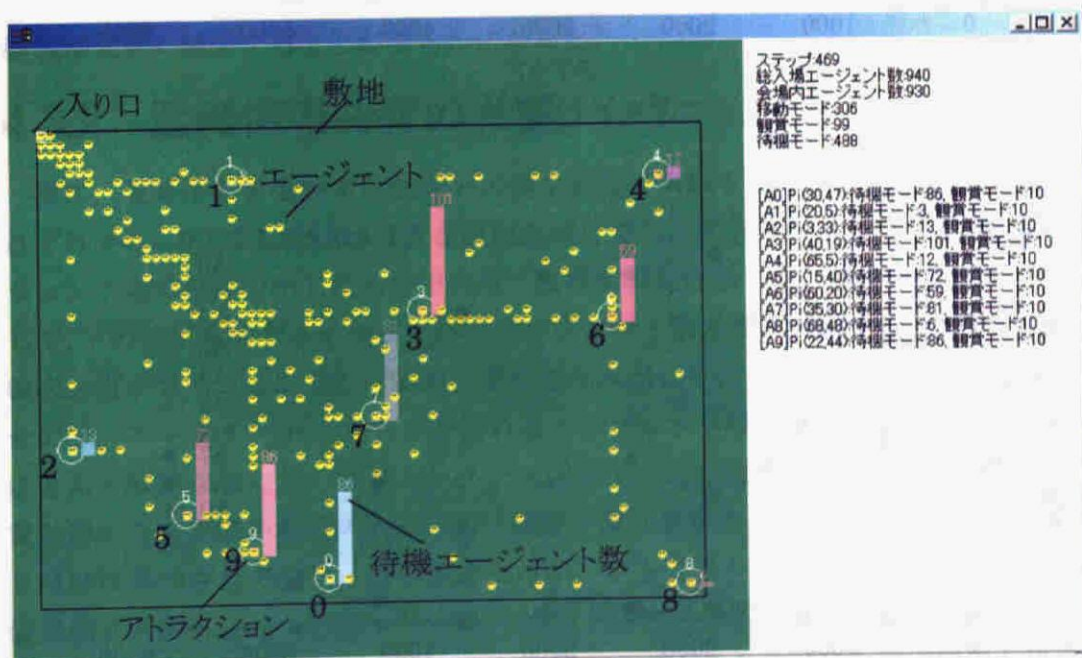


図4.1：シミュレーション実行画面

はじめに、施設配置の違いのみが混雑に与える影響について考える。図4.2は、総入場エージェント数、現在会場に存在するエージェント数およびエージェントのモードの推移を30ステップごとに示したものである（以後、各図および結果は30ステップごとのデータに基づくものとする）。総入場エージェント数とはそのステップまでに入場したエージェントの総数である。同図より、初期状態から会場内のエージェント数が増加し移動モードの来場者が大きく増加していくのに次いで、アトラクションに到達して観賞モードに入るエージェントが大きく増加していく様子がわかる。観賞モードのエージェント数が最大値100（10人×10カ所）付近に達すると、今度は待機モードのエージェントが増加する。その後、会場内のエージェント数が

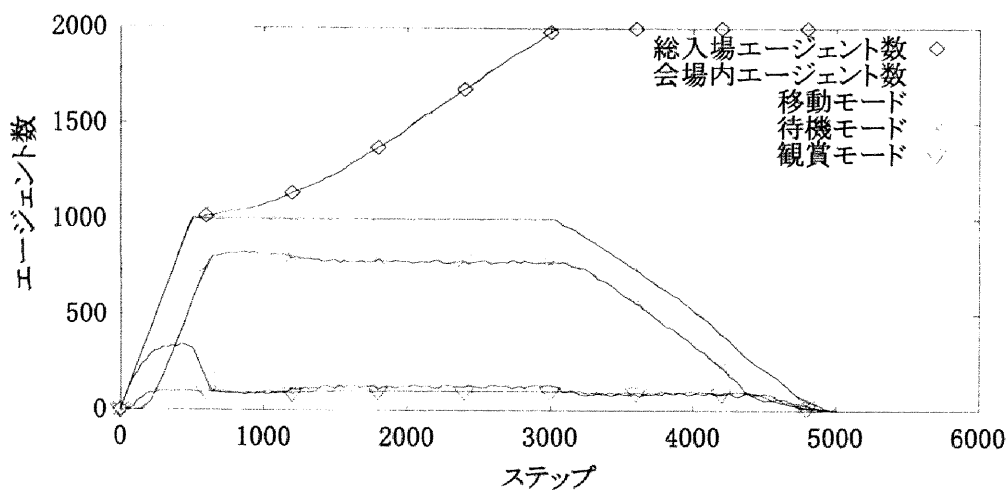


図4.2：エージェントのモードの推移 ($r=0.0$)

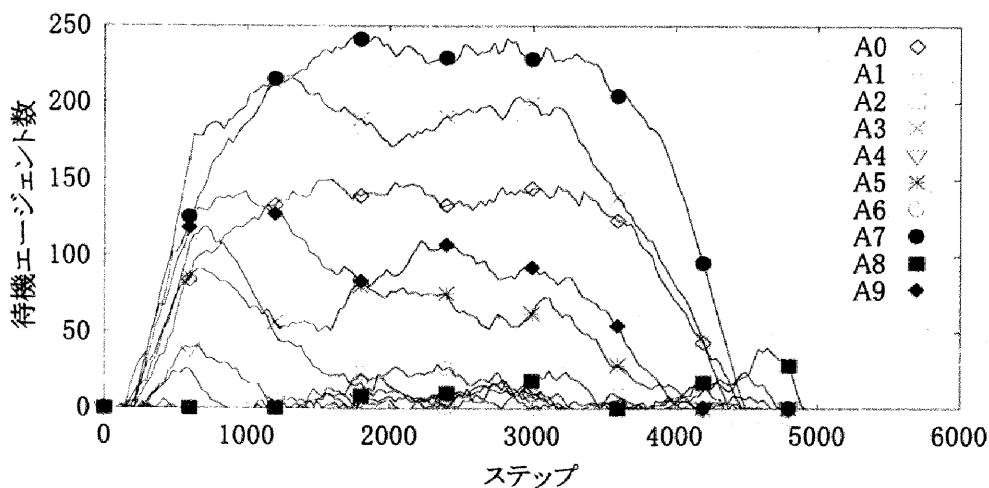


図4.3：各アトラクションの待機エージェント数の推移 ($r=0.0$)

1000 に達すると、エージェントの入場制限によって会場内のエージェント数が頭打ちとなるが、退場するエージェント数だけさらにエージェントが入場するため、総入場エージェント数は徐々に増加する。しばらくこのような状態が続いた後、退場に伴う来場者の減少により待機モードのエージェント数が徐々に減少していき、最後にはすべてのエージェントが退場した。図 4.3 は、混雑情報を全く提供しない場合（提供確率 $r=0.0$ ）において、アトラクションごとの待機エージェント数の推移を 30 ステップごとに示したものである（グラフの番号は図 4.1 のアトラクションにつけられた番号に対応する）。本研究では、このアトラクションごとの待機エージェント数の偏りを、集団における行動の多様性を示すものとして注目する。同図より、待機エージェント数のアトラクションごとの推移には大きな違いがあり、全体として会場の中心（アトラクションの重心）付近に位置するアトラクション 3, 7 などが混雑しやすい傾向があることがわかる。逆に、離れた位置にあるアトラクション 1 や 8 では、シミュレーション序盤・中盤において待機エージェントが存在せず、収容可能な人数をも満たさない状態が断続的に続いており、効率的でない状況であると言える。また、これらのアトラクションでは、待機エージェント数がシミュレーションの終盤で急増する傾向があった。これは、他のアトラクションから遠い位置にあることが、エージェントが観賞するのを後回しにさせ、終盤において多くのエージェントが集中したためであると考えられる。このように、人気度や収容能力・観賞時間が同一であっても、その配置によって混雑の様子は大きく異なり、来場者がいくつかのアトラクションに集中してしまうという意味で、エージェントの行動の多様性は大きく減少することが明らかになった。

4.3.2 混雑情報共有の影響

次に、混雑情報を提供する確率 r の値を変えて実験を行った。図 4.4~4.9 は、それぞれ $r=0.1, 0.2, 0.3, 0.5, 0.8, 1.0$ を用いた試行での、アトラクションごとの待機エージェント数の推移を示したものである。混雑情報を提供しない場合（図 4.3）と比較して、 r の値が増加するに従って待機エージェント数の偏りが減少しており、混雑情報の共有がエージェントの行動の多様性を生み出していることがわかる。 $r=0.2$ の場合では、アトラクション 1 や 8 を訪れるエージェント数が収容能力を下回る状況はほとんど解消されたが、各アトラクションをすべての観賞予定のエージェントが観賞し終わるまでにかかるステップ数には若干偏りがある。 $r=0.3$ の場合では、この偏りはほぼ解消されており、効率的であるが、アトラクションごとの行列の長さには依然偏りがある。

$r=0.5$ の場合では、細かく振動しながらもどのアトラクションもほぼ同様な待機エージェント数で推移し、シミュレーション中盤において最も待機エージェント数が多くなるような緩やかな曲線を描いており、待機エージェント数の偏りが大きく緩

和されていることがわかる。ところが、 $r=0.8$ の場合や常に混雑情報を提供する $r=1.0$ の場合では、逆に振動が激しくなっていることがわかる。この振動は、エージェントにとって行列に並ぶタイミングによって不公平を生むという点で、望ましいものとは言えない。そこで、 r の値の変化に伴ってアトラクション毎の待機エージェント数の偏りがどのように変化するかを定量的に明らかにするために、 $r=0.0\sim 1.0$ まで値を 0.1 ずつ変えて実験を行ったときの、全ステップにおけるアトラクションごとの待機エージェント数の標準偏差の平均を示したのが図 4.10 である。各値は 10 試行の平均である。 r が大きくなるに従って標準偏差が大きくなり減少し、 $r=0.5$ で最小となり、その後緩やかに増加していることがわかる。これは、この条件では $r=0.5$ が、行動の多様性が最大という意味において、適当な情報提供・共有の頻度であることを示している。

r が一定の値を超えると振動が発生する理由は、アトラクションの収容能力が有限であることと、各混雑情報に従った結果が反映されるまでに時間差があるという意味で、エージェントが時間遅れを伴う情報に従っていることにある。次のような状況を考える。いま、あるアトラクション A が混雑の最も少ない状況になったと仮定する。このとき、他のアトラクションを観賞し終えたエージェントは、混雑情報からアトラクション A を目指すようになる。しかし、これらのエージェントはアトラクション A に実際に到達するまでには移動時間がかかるために、実際に行列が長くなるまでには時間遅れが存在する。その間も情報端末はアトラクション A を目指すべきものとして混雑情報を提供し続けるために、この間にあまりに多くのエージェントが混雑情報の提供を受けて情報を共有すると、過剰な数のエージェントがアトラクション A に向かってしまうことになる。このようなことが連続して起きることで、上記のような混雑の振動が生じていると考えられる。

また、試行を重ねた結果、配置の影響で混雑が少ない傾向にあったアトラクション 1, 4, 8 などにおいて、この振動が大きい傾向が観察された。これは、他のアトラクションからこれらのアトラクションに到達するまでに長い時間がかかるために、前述の時間遅れの効果が大きいことが原因であると考えられる。混雑情報の提供とその共有は、混み合わない傾向にあるアトラクションへエージェントを分散させることでエージェントの行動に多様性をもたらしていたが、その頻度が高すぎる場合には、逆に混み合わない傾向にあるアトラクションへの過度の集中化を引き起こしていると考えられる。

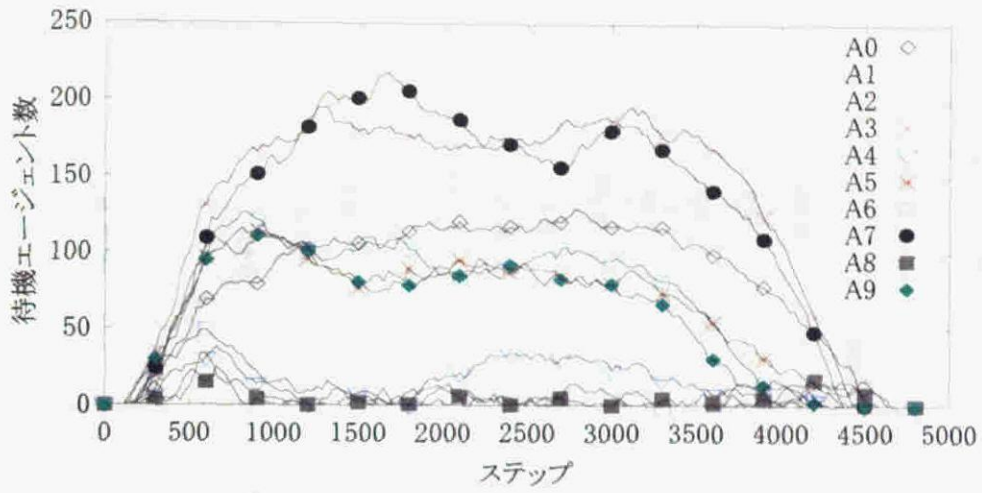


図 4.4 : 各アトラクションの待機エージェント数の推移 ($r=0.1$)

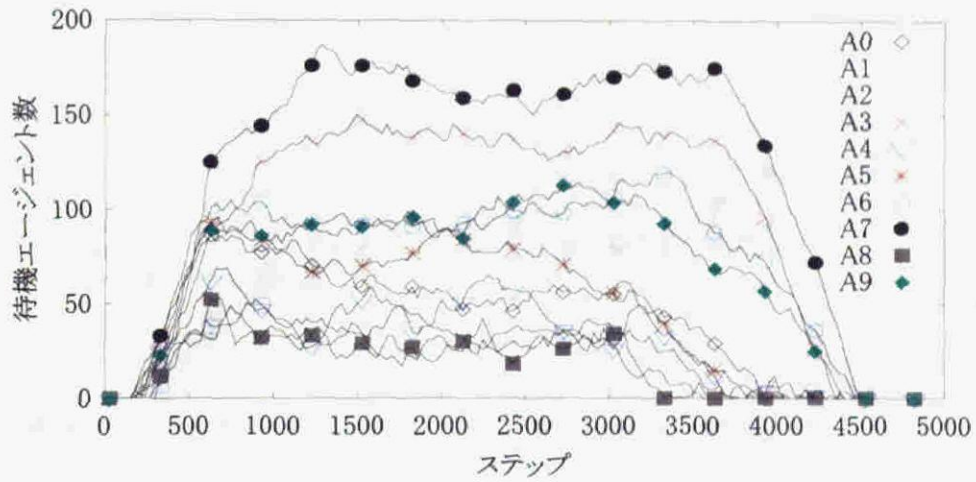


図 4.5 : 各アトラクションの待機エージェント数の推移 ($r=0.2$)

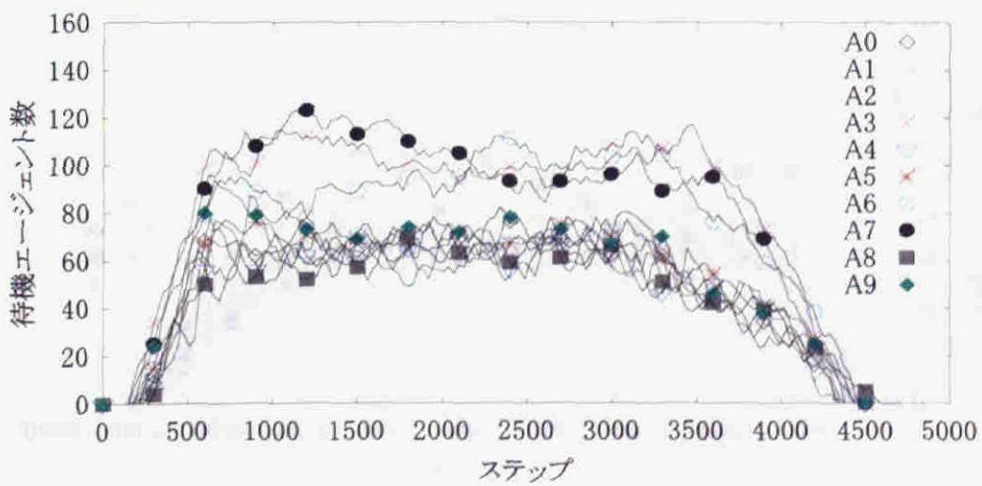


図 4.6 : 各アトラクションの待機エージェント数の推移 ($r=0.3$)

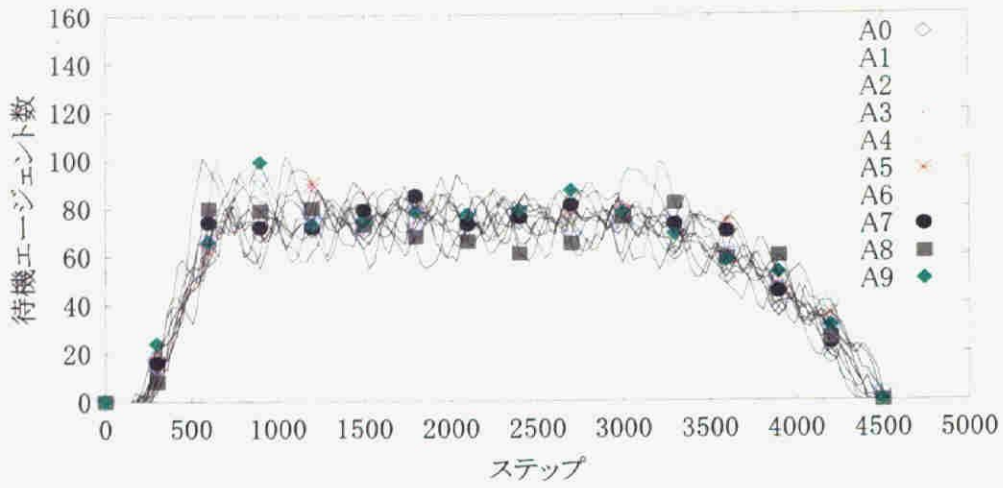


図 4.7：各アトラクションの待機エージェント数の推移 ($r=0.5$)

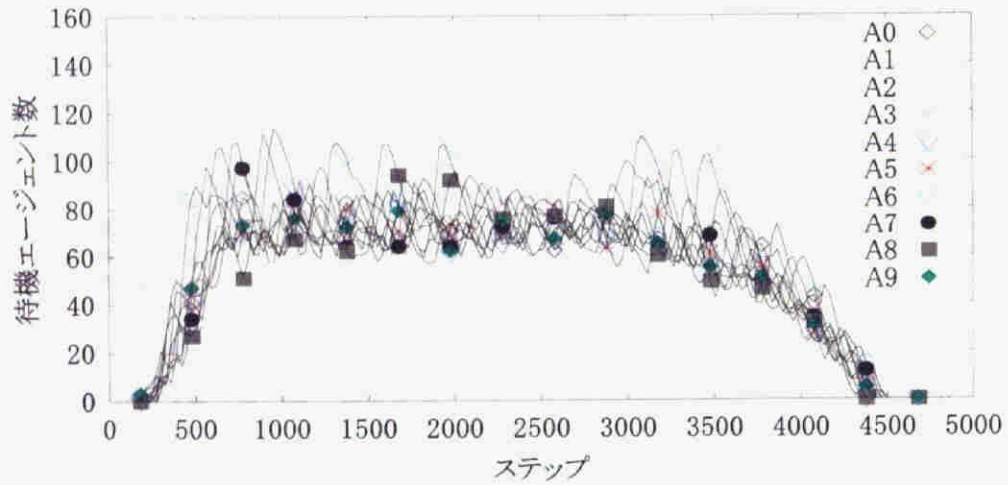


図 4.8：各アトラクションの待機エージェント数の推移 ($r=0.8$)

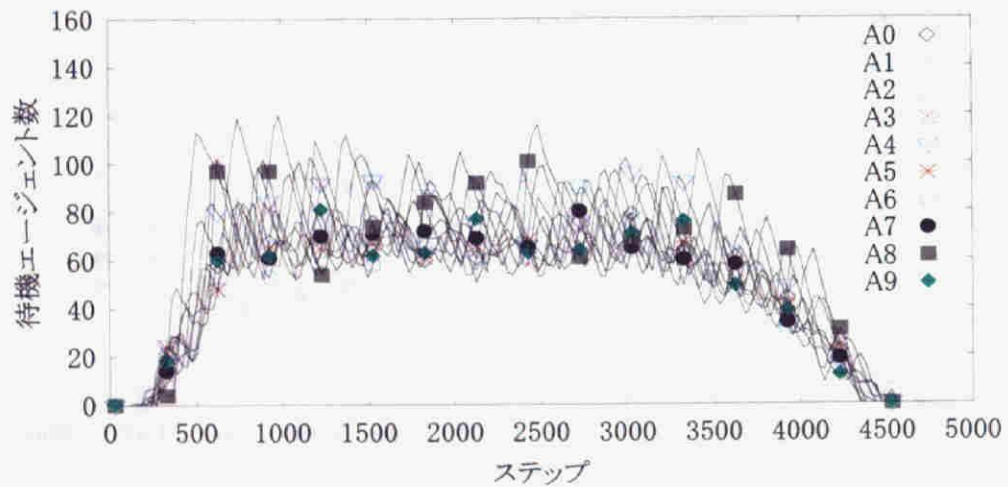


図 4.9：各アトラクションの待機エージェント数の推移 ($r=1.0$)

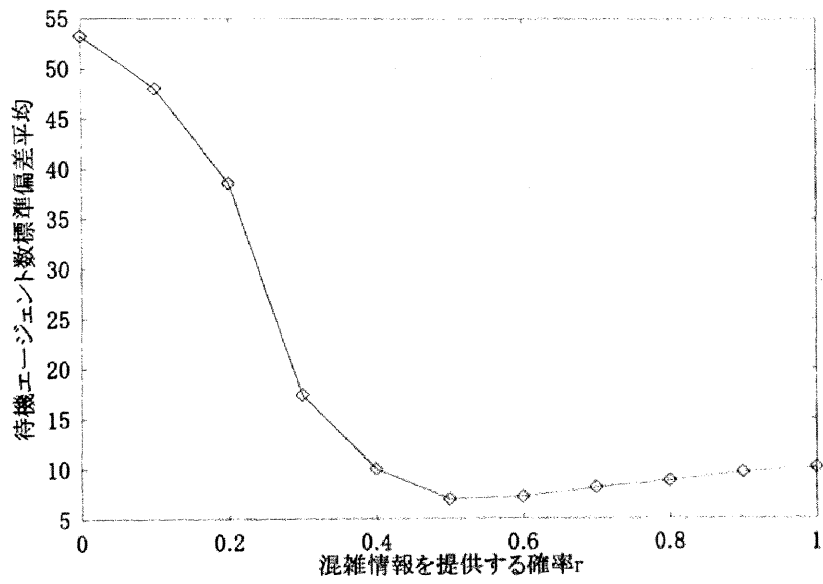


図 4.10 : 混雑情報提供確率 r と待機エージェント数の標準偏差

4.4 情報共有の効果の変動と行動多様性

4.4.1 「突発待ち行列」の発生

これまでの実験で、アトラクションごとの待機エージェント数の偏りを最小にするような混雑情報提供・共有確率が存在し、それは $r=0.5$ の場合において最小となることがわかった。この $r=0.5$ という値自体は実験の設定に依存するものであるが、例えば入場者数のみを大きく増やし、より長い期間試行を行った場合においても、偏りが最小である挙動が観察され続けるだろうか。そこで、これまでの設定において、全エージェント数 N_p のみをこれまでの10倍の20000人に設定して $r=0.5$ で実験を行った（前節の値と比較するとやや極端な設定であるが、本節ではエージェント数の増加によるシミュレーション時間の増加がもたらす影響をより明確に議論するために、この条件を採用した）。その結果、いくつかの試行で次のような注目すべき現象が確認された。

図4.11は、ある試行でのアトラクション毎の待機エージェント数の推移を示したものである。試行の初期では、これまでの結果と同様に細かな振動を伴いながらのアトラクションの待機エージェント数もほぼ同様の推移を保っているが、約15000ステップから20000ステップ、25000ステップ以降で突如アトラクション7の待機エージェント数のみが大きく増加していることがわかる。本研究では、このように、初期状態からしばらくの間どのアトラクションの待機エージェント数も同様の値で推移したのち、ある一つのアトラクションの待機エージェント数のみが突発的に大きく増加する現象を、「突発待ち行列」の発生と呼ぶことにする。試行を繰り返した

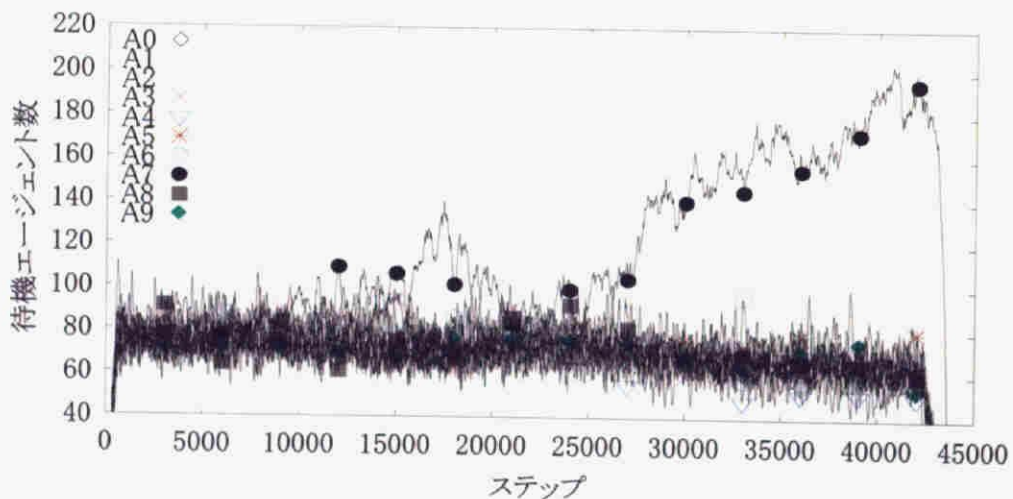


図4.11：各アトラクションの待機エージェント数の推移 ($r=0.5, N_p=20000$)

結果、突発待ち行列が発生するアトラクションは、混雑情報の提供なしの場合において配置の影響により混雑しやすいアトラクションであることが判明した。

4.4.2 情報共有による行動多様性維持の効果の減少

このような現象が起きる理由は、混雑情報共有によるエージェントの分散化、すなわち、行動の多様性維持の効果が、シミュレーションが進行するにつれて減少してくるためであると考えられる。これは、エージェントの将来の行動規則の多様性が減少するという、情報共有の副作用によって引き起こされると考えられる。

これまでの議論からわかるように、混雑情報は、それを受け取ったエージェントに対し、配置の影響によって混雑しない傾向にあるアトラクションを優先的に観賞させる働きがある。このため、エージェントは混雑情報に従うことを繰り返すほど観賞予定のアトラクションの内に混雑しにくいアトラクションが少なくなる傾向がある。図 4.12 は、図 4.11 と同一の試行において、各アトラクションについて、それを観賞予定のアトラクションに含んでいる会場内のエージェントの数（観賞予定エージェント数）の推移を示したものである。これらの偏りは、会場内のエージェントの将来の行動の多様性を象徴するものと捉えることができる。

同図より、シミュレーションが進行するに従って観賞予定エージェント数はアトラクションごとに偏りが生じ、突発待ち行列が発生したアトラクション 7 の観賞予定エージェント数は相対的に大きな値となっていることがわかる。これは、混雑情報共有によって、配置の影響で混雑する傾向にあるアトラクションを後回しにする効果が実際に累積されていることを示している。これが、即時の行動の多様性が増加する一方で、将来の行動規則の多様性が減少するという情報共有の副作用である。

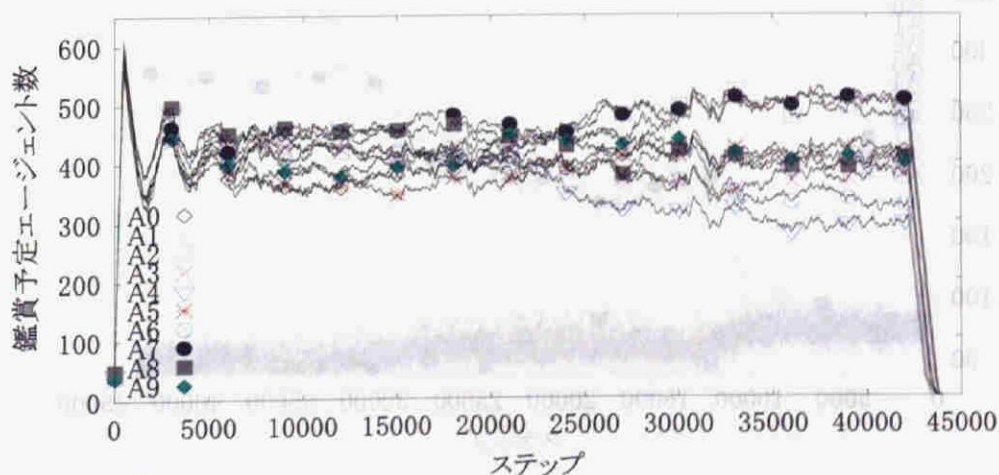


図 4.12 : 各アトラクションの観賞予定エージェント数の推移 ($r=0.5$, $N_T=20000$)

その結果、情報を受け取っても混雑しやすいアトラクションしか選ぶことができず、そこに行かざるを得ないエージェントが徐々に出現する。

$N_p=2000$ の場合はこの影響が表出する前にシミュレーションが終了したが、 $N_p=20000$ の場合は、相対的な差が大きく広がった結果、情報を共有しても混雑する傾向にあるアトラクションに向かうエージェントが増加し、そのアトラクションの待機エージェント数が増加を始めたと考えられる。更に、一旦あるアトラクションにおける行列が長くなると、新たに入場したエージェントは混雑の少ないアトラクションをすばやく巡回できてしまうために、混雑しているアトラクションはより混雑し、それ以外のアトラクションはより混雑しにくくなるという正のフィードバック効果が働き、突発待ち行列の発生を招いたと考えられる。

以上のような時間経過のシナリオが考えられるものの、図 4.11, 4.12 を比較すると、観賞予定エージェント数の偏りと突発待ち行列の発生には強い関連があることは分かるが、その因果関係はやや明らかでない面がある。そこで、これらの因果関係を明確にするために、シミュレーションの進行途中の 10000 ステップ目において、会場内に存在するエージェントのうち、アトラクション 7 を観賞予定としていない 100 個体のエージェントをランダムに選び出し、その個体の観賞予定のアトラクションにアトラクション 7 を追加するという操作を加えて実験を行った。図 4.13, 4.14 はこの操作を追加した時の、各アトラクションの待機エージェント数、観賞予定エージェント数の推移を示したものである。図 4.14 から、上記の操作により 10000 ステップにおいてアトラクション 7 の観賞予定エージェント数が大きく増加していることがわかる。同時に、図 4.13 より、10000 ステップを過ぎた直後からアトラクション 7 の待機エージェント数が急激に増加し、突発待ち行列が発生していることが

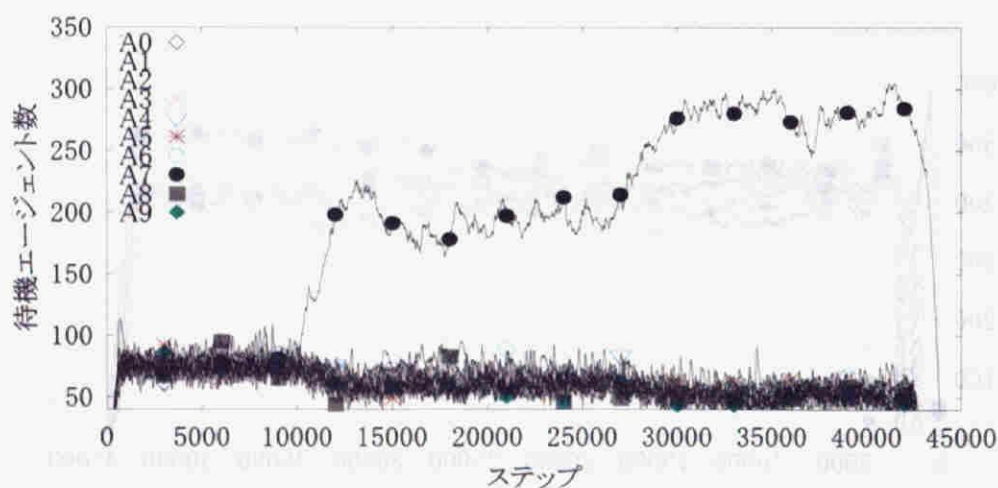


図 4.13 : 10000 ステップ目で操作を加えた場合の待機エージェント数の推移
($r=0.5$, $N_p=20000$)

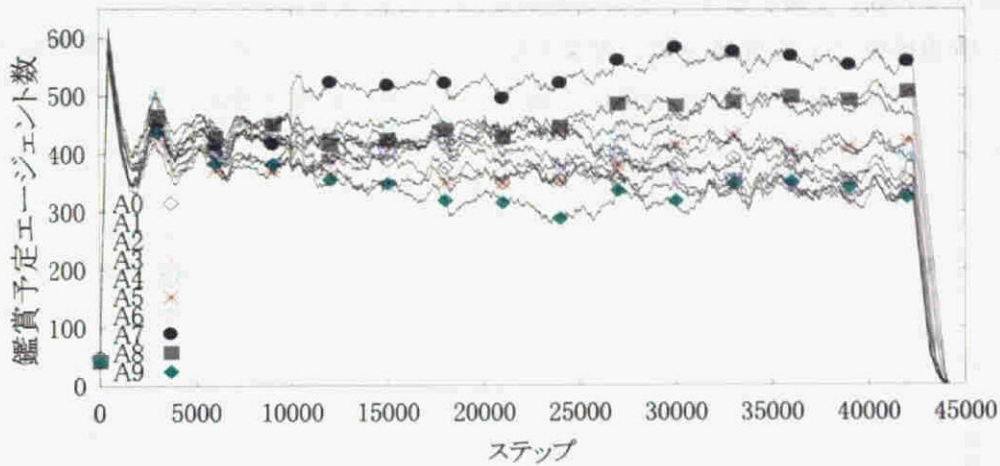


図 4.14 : 10000 ステップ目で操作を加えた場合の観賞予定エージェント数の推移
($r=0.5$, $N_p=20000$)

わかる。更なる実験の結果、この操作を加えることにより、ほとんどすべての試行においてアトラクション 7 で突発待ち行列が発生することが判明した。また、この操作を、混雑情報の提供なしの場合において混雑しにくかったアトラクション 8 に対して行った実験では、アトラクション 8 に突発待ち行列は発生しなかった。以上から、混雑情報の提供なしの場合において混雑しやすいアトラクションにおける観賞予定エージェント数の偏りの増加は、そのアトラクションの突発待ち行列の発生の原因となっていることが明らかになった。

4.4.3 突発待ち行列の定量的解析

最後に、混雑情報を提供する確率 r と突発待ち行列が発生する頻度、および、系全体の効率との関係を定量的に解析する。そのために、この現象を、「初期状態から 12000 ステップまでの間において、すべてのアトラクションにおける待機エージェント数が常に 120 以下で推移し、かつ、それ以降のステップからシミュレーション終了までの間において、ただ一つのアトラクションの待機エージェント数が 160 以上であり、それ以外のアトラクションではすべて 100 以下であるステップが少なくとも一つ存在する試行」と定義する。

図 4.15 は $N_p=20000$ の設定で $r=0.0\sim 1.0$ まで値を 0.1 ずつ変えてそれぞれ 30 回試行を行ったとき、各設定において上記の条件を満たした試行の回数と、各設定での全ステップにおけるアトラクションごとの待機エージェント数の標準偏差の平均を表したものである。同図より、上記のような現象は $r=0.4\sim 0.7$ の間で観察され、 $r=0.5$ の場合においてその頻度が最も高いことがわかる。会場全体の効率という観点からは、突発待ち行列の発生は好ましいことではないが、 r に関してこのような動的現象が生じる特殊な条件が存在するのは興味深いと言える。また、待機エージェント数

の標準偏差は $r=0.7$ で最小値をとることがわかる。このことは、 r が中程度の条件において、情報提供による混雑分散の効果の低下により上記の現象が発生し標準偏差が増加するために、前節での $N_p=2000$ の場合よりも高い頻度が最適値となることを示している。

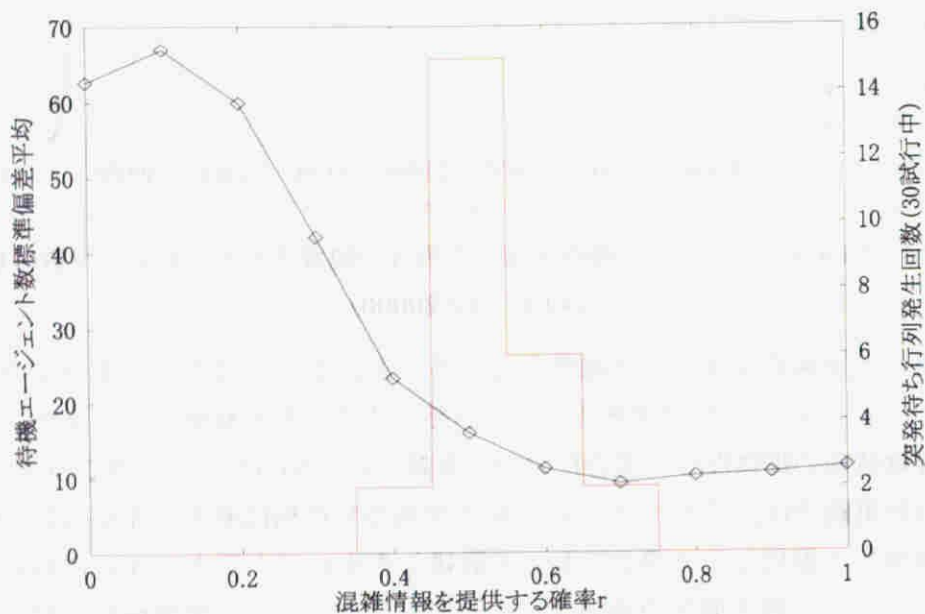


図 4.15：混雑情報提供確率 r と待機エージェント数標準偏差，および、突発待ち行列発生回数（30 試行， $r=0.5$ ， $N_p=20000$ ）

4.5 まとめ

本章では、イベント会場における来場者の作る混雑に施設配置と混雑情報の提供がどのような影響を及ぼすかについて、集団における行動の多様性に対する情報共有の影響とその適応性に関する具体的な状況設定として議論してきた。

はじめに、来場者が観賞予定のアトラクションを近くのものから順に観賞するという条件では、アトラクションの配置に特有な混雑の偏りの推移を示すことが明らかになった。次に、来場者に携帯情報端末を用いて混雑状況を提供し、情報の提供を受けたエージェントはそれに従って混雑の少ないアトラクションを観賞する設定で実験を行った。このとき、情報を提供する頻度に様々な条件を用いて実験を行ったところ、混雑情報の提供は基本的にはエージェントの行動の多様性を増加させ、アトラクション毎の待機エージェント数の偏りを減少させるが、頻度が高すぎるとかえって行動の多様性を減少させ、待機エージェント数の振動が大きくなるという影響があることが判明した。

更に、来場者数の少ない設定で混雑の偏りが最小となる情報提供の頻度を用いても、来場者を増やして長い期間実験を行うと、時間経過に伴って混雑情報提供の効果が減少することにより将来の行動多様性が減少し、それをきっかけにして突発的に長い待ち行列が発生することが判明した。また、この条件では、将来の行動多様性の減少の影響が現れない、より高い情報提供の頻度が、混雑の偏りを最小にすることも明らかになった。この突発待ち行列の発生は、エージェントが共有する情報の多様性と実際に行われる行動の多様性という2つの多様性間で生じる影響が、時間経過と共に変動する現象を示したという点で、興味深いと言える。

最後に、これらの議論で得られた知見は、このようなエージェントベースモデリングに基づくアプローチが情報化社会における情報提供機器・システムの設計の新しい方法論になりうることを示すことについて述べる。携帯情報端末や情報提供システムにおいて、各個人が受け取ったり共有したりした情報に基づいて行動することが集団全体の行動にどのような振る舞いをもたらすかは、その性能評価のために極めて重要な点である。にもかかわらず、そのような機器やシステムを設計する段階では、人間の集団行動に対して実際に与える影響について、その理解は容易ではないと言える。本研究で得られた成果は、このような問題に対し、情報提供システムとそれに基づいて行動する集団をモデルに同時に盛り込むことにより、情報提供と集団全体の行動との間に生じるダイナミクスを明らかにすることが可能であることを示唆するものである。

第5章 結論

本研究では、個体間相互作用に基づく創発的ダイナミクスについて知見を得るために、3つの具体的な研究テーマを取り上げて議論してきた。はじめに、進化と学習という2つの異なる適応プロセス間の相互作用を明らかにするため、表現型可塑性進化を導入した繰り返し囚人のジレンマゲームにおける戦略の進化モデルを用いて実験を行った。その結果、進化の過程で、一旦学習に依存した協調集団へと進化したのち、学習によるコストのために学習に依存しない協調集団へと進化する Baldwin 効果が確認された。また、最終的に集団中を占めた戦略は、必要最小限の可塑性をうまく利用した極めて興味深い特徴を備えるものであった。これらから、本モデルでは、2つの適応プロセスは系全体の協調集団への進化に対して協調的に働いたと言える。また、このような2つの適応プロセス間の相互作用に関する議論を行う上では、本研究で採用したような各プロセスへの集団の依存度が自動的に調節されるようなメカニズムの存在が重要であることが示された。

次に、個体間における相互作用の枠組みの違いやそれ自体の進化がどのような影響を及ぼすかについて理解するために、空間的局所性とその進化を導入したN人版繰り返し囚人のジレンマゲームにおける戦略の進化モデルを構築し、進化実験を行った。まず、従来同一のものとして議論されていた、相互作用に関する局所性と、世代交代に関する局所性を区別して実験を行い、環境条件としての相互の影響について解析した。その結果、両者は協調行動の創発に大きく異なる影響を与えることが明らかになった。次に、それら自身が個体の戦略の一部として進化する条件で実験を行ったところ、相互作用に関する局所性は世代を通して増加減少を繰り返し、世代交代に関する局所性はそれにかかるコストに依存してある値に収束するという異なる進化ダイナミクスが確認された。また、相互作用に関する局所性の動的な進化は、その進化なしの場合と比較して、協調行動の創発を促進する場合があることが判明した。これらの結果は、相互作用の枠組みが系全体の挙動を大きく左右することを示すと同時に、枠組み自体が変動するようなメカニズムの存在が、系自体をさらに適応的な状態へと導く可能性を示すものである。

最後に、集団内で共有される情報の多様性が、個体が行う行動の多様性と系全体の挙動に与える影響について明らかにするために、イベント会場における来場者の作る混雑に施設配置と混雑情報の提供がどのような影響を及ぼすかについて論じた。その結果、来場者に提供する混雑情報提供の頻度を調節することで、施設配置によって生じる来場者の混雑の偏りを緩和できることが明らかになった。しかし、同じ頻度の条件であっても、エージェント数を増加して長期間実験を行うと、情報共有のもう一つの効果である、将来の行動規則の多様性の減少によって、試行の途中で突然混雑が発生することが判明した。これは、エージェントが共有する情報の多様

性と実際に行われる行動の多様性という 2 つの多様性の間に生じる影響が、時間経過と共に変動するというシナリオを示したという点で、興味深いと言える。

以上、本研究は、エージェントベースモデリングに基づく 3 つの研究事例を通して、個体間相互作用に基づく創発的ダイナミクスについて議論してきた。最後に、序論で掲げた 3 つの本質的問いに関して得られた知見について改めてまとめると次のとおりになる。

- 1) 系を駆動する複数の適応プロセスは、系全体をより適応的な状態へともたらず方向へと協調的に駆動しうる。その過程において、各適応プロセスが系をどれだけ支配しているかが変化しうることは重要な点となる。
- 2) 相互作用の構造は系全体の挙動に大きな影響を与えうるが、その影響の仕方は相互作用の種類によって異なる。また、相互作用の構造自体が適応プロセスのような内的要因により変動し続けることが、それが外的要因により固定された状況では観察できないような挙動を系にもたらしうる。
- 3) 異なるレベルの多様性は相互に影響しうる。例えば、あるレベルの多様性が、もう一方の多様性を維持するなどの影響を与えうる。しかし、その影響の仕方はかならずしも定常的ではなく、突如としてその影響が小さくなってしまふといった動的な変化が起こりうる。

以上の知見を総合して、個体間相互作用に基づく創発的ダイナミクスに関して言えることは、系に創発的もしくは適応的な状態が生じるには、ある種のやわらかさが必要とされるということである。しかし、それがどこに存在するかは、これまで示したとおり、複数の適応プロセスへの依存度の変化であったり、相互作用の構造の変化であったり、多様性の変動であったりと、文脈によって様々である。したがって、系に生ずる創発的ダイナミクスを理解する上では、どこに、どのような形で存在するやわらかさが創発性をもたらしているかを明らかにすることが重要な鍵になると言えよう。

エージェントベースモデリングに基づく構成論的研究においては、その多くが本研究で採用してきたような抽象的なモデルを用いた議論がなされている。そのため、個々の具体的状況設定の理解に留まらず、同様なダイナミクスを持つ系について適用可能で普遍的な議論を展開できる可能性がある。本研究においても、それぞれの事例をもとに、創発的ダイナミクスに関して上記のような興味深い知見を得ることができた。今後も様々な状況設定に基づくモデル化と、そこに生じる現象の理解を重ねることで、創発的ダイナミクスに関してより本質的な知見を得ることが期待される。

謝辞

本研究は、名古屋大学大学院人間情報学研究科において、名古屋大学大学院人間情報学研究科助教授有田隆也先生のご指導のもとで行われたものであります。有田隆也先生には、大学院入学当初の右も左もわからない頃から、人工生命研究に対する姿勢はもとより、研究の進め方、議論の仕方、論文の書き方など何から何までご教授いただきました。先生の熱心なご指導なくして、このように研究をまとめることは到底できませんでした。ここに、心より感謝の意を表します。

研究を進めていく中、少ない機会にもかかわらず、貴重なご助言をいただきました。電気通信大学電気通信学部講師佐山弘樹先生に深く感謝致します。

また、本研究を円滑に進めるにあたり、研究助成をしてくださいました堀情報科学振興財団に深く感謝致します。

公私にわたり支援して下さった、名古屋大学大学院人間情報学研究科有田研究室（人工生命ラボトリエ）の諸氏に感謝致します。特に、博士後期課程の赤石仁、中道義之の両氏には、日頃から充実した議論を交わすだけでなく、同じく研究者を目指す者として何度となく勇気づけられました。ここに感謝致します。

最後に、これまで暖かく見守って下さった両親と妹に心より感謝致します。

2003年1月

参考文献

- [Ackley 91] Ackley, D., Littman, M.: Interaction between Learning and Evolution, *Artificial Life II*, pp. 487-509, Addison-Wesley (1991).
- [Akaishi 02] Akaishi, J. and Arita, T.: Misperception, Communication and Diversity, *Proceedings of Artificial Life VIII*, pp. 350-357 (2002).
- [Anderson 95] Anderson, R. W.: Learning and Evolution: A Quantitative Genetics Approach, *Journal of Theoretical Biology*, Vol. 175, pp. 89-101 (1995).
- [有田 00] 有田隆也: 動的な適応度地形の概念に基づく生命性に対する構成論的アプローチ, *科学哲学*, 33-2 (2000).
- [Arita 00] Arita, T. and Suzuki, R.: Interactions between Learning and Evolution: The Outstanding Strategy Generated by the Baldwin Effect, *Proceedings of Artificial Life VII*, pp. 196-205 (2000).
- [有田 02] 有田隆也: 人工生命, 医学出版 (2002).
- [Arthur 94] Arthur, W. B.: Inductive Reasoning and Bounded Rationality, *American Economic Association Papers Proceedings*, Vol.84, No.2, pp.406-411 (1994).
- [Axelrod 84] Axelrod, R.: *The Evolution of Cooperation*, Basic Books (1984).
- [Banks 94] Banks, T.: Exploring the Foundations of Artificial Societies: Experiments in Evolving Solutions to Iterated N-player Prisoner's Dilemma, *Proceedings of the Fourth International Conference on Artificial Life*, pp. 337-342 (1994).
- [Baldwin 1896] Baldwin J. M.: A New Factor in Evolution, *American Naturalist*, Vol. 30, pp. 441-451 (1896).
- [Boerlijst 97] Boerlijst, M. C., Nowak, M. A. and Sigmund, K.: The Logic of Contrition, *Journal of Theoretical Biology*, Vol. 185, pp. 281-293 (1997).
- [Bonacich 76] Bonacich, P., Shure, G. H., Kahan, J. P. and Meeker, R. J.: Cooperation and Group Size in the N-Person Prisoners' Dilemma, *Journal of Conflict Resolution*, Vol. 20, No. 4, pp. 687-706 (1976).
- [Challet 97] Challet, D. and Zhang, Y. C., Emergence of Cooperation and Organization in an Evolutionary Game, *Physica A*, vol.246, pp.407-418 (1997).
- [Grim 96] Grim, P.: Spatialization and Greater Generosity in the Stochastic Prisoner's Dilemma, *Bio Systems*, Vol. 37, pp. 3-17 (1996).
- [Harvey 93] Harvey, I.: The Puzzle of the Persistent Question Marks: A Case Study of Genetic Drift, *Genetic Algorithms: Proceedings of the Fifth International Conference*, pp. 15-22 (1993).
- [辺見 02] 辺見和晃: 来場者に優しいテーマパークー混雑緩和問題と情報の共有, 山影 進, 服部正太編, コンピュータの中の人工社会, pp. 124-139 (2002).

- [Hinton 87] Hinton, G. E. and Nowlan, S. J.: How Learning Can Guide Evolution, *Complex Systems*, Vol. 1, pp. 495-502 (1987).
- [Hogg 91] Hogg, T. and Huberman, B. A.: Controlling Chaos in Distributed Systems, *IEEE Transactions on Systems, Man and Cybernetics*, Vol.21, No.6, pp.1325-1332 (1991).
- [Ishibuchi 99] Ishibuchi, H., Nakari, T. and Nakashima, T.: Evolution of Neighborly Relations in a Spatial IPD Game with Cooperative Players and Hostile Players, *Proceedings of Congress on Evolutionary Computation*, pp.929-936 (1999).
- [石渕 00] 石渕久生, 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イヤー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).
- [Joshi 87] Joshi, N. V.: Evolution of Cooperation by Reciprocation within Structured Demes, *Journal of Genetics*, Vol. 66, No. 1, pp. 69-84 (1987).
- [Kaneko 00] Kaneko, K. and Yomo, T.: Sympatric Speciation from Interaction-induced Phenotypic Differentiation, *Proceedings of Artificial Life VII*, pp. 113-121 (2000).
- [Kauffman 93] Kauffman, S. A.: *The Origins of Order: Self Organization and Selection in Evolution*, Oxford University Press (1993).
- [Lindgren 91] Lindgren, K.: Evolutionary Phenomena in Simple Dynamics, *Artificial Life II*, pp. 295-311 (1991).
- [Lindgren 94] Lindgren, K. and Nordahl, M. G.: Evolutionary Dynamics of Spatial Games, *Physica D*, Vol. 75, pp. 292--309 (1994).
- [Matsushima 98] Matsushima, M. and Ikegami, T.: Evolution of Strategies in the Three-person Iterated Prisoner's Dilemma Game, *Journal of Theoretical Biology*, Vol. 195, pp. 53-67 (1998).
- [Mayley 96] Mayley, G.: Landscapes, learning costs and genetic assimilation, *Evolutionary Computation*, Vol. 4, No. 3, pp. 213-234 (1996).
- [Maynard-Smith 82] Maynard-Smith, J.: *Evolution and the Theory of Games*, Cambridge University Press (1982).
- [Nowak 92a] Nowak, M. A. and May, R. M.: Evolutionary Games and Spatial Chaos. *Nature*, Vol. 359, pp. 826-829 (1992).
- [Nowak 92b] Nowak, M. A. and Sigmund, K.: Tit for Tat in Heterogeneous Populations, *Nature*, Vol. 355, pp. 250-252 (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, No. 1, pp. 56-58 (1993).
- [Pigliucci 01] Pigliucci, M.: *Phenotypic Plasticity*, The John Hopkins University Press

- (2001).
- [Sasaki 99] Sasaki, T. and Tokoro, M.: Adaptability of Darwinian and Lamarckian Populations toward an Unknown New World, *Proceedings of the 5th European Conference on Artificial Life*, pp. 39-48 (1999).
- [Savit 99] Savit, R. Manuca, R. and Riolo, R.: Adaptive Competition, Market Efficiency, and Phase Transitions, *Physical Review Letter*, Vol.82, No.10, pp.2203-2206 (1999).
- [Seo 99] Seo, Y. G., Cho, S. B. and Yao, X.: Emergence of Cooperative Coalition in NIPD Game with Localization of Interaction and Learning, *Proceedings of the 1999 Congress on Evolutionary Computation*, Vol. 2, pp. 877-884 (1999).
- [鈴木 99a] 鈴木麗聖: 囚人のジレンマゲームにおける Baldwin 効果, ALIREN (人工生命研究会) +ALIST (在京若手研究者による人工生命自主セミナー) 合同研究会ポジションペーパー (1999).
- [鈴木 99b] 鈴木麗聖, 有田隆也: 囚人のジレンマゲームにおける Baldwin 効果, 人工知能学会第 13 回全国大会論文集, pp. 277-278 (1999).
- [鈴木 99c] 鈴木麗聖, 有田隆也: メタ・パブロフ: 進化と学習による適応を自動調節する繰り返し型囚人のジレンマ戦略, 情報処理学会研究報告, 99-GI-1, pp. 15-22 (1999).
- [鈴木 99d] 鈴木麗聖, 有田隆也: 学習が進化の過程に与える影響に関するシミュレーション解析, 第 9 回数理生物学シンポジウム予稿集 (数理生物学懇談会ニュースレター), p. 9 (1999).
- [鈴木 00a] 鈴木麗聖, 有田隆也: 進化と学習の相互作用 - 繰り返し囚人のジレンマゲームにおける Baldwin 効果 -, 人工知能学会誌, Vol. 15, No. 3, pp. 495-502 (2000).
- [鈴木 00b] 鈴木麗聖: 動的環境における進化と学習の相互作用～Baldwin 効果の工学的応用に向けて～, 第 1 回 MYCOM (AI 若手の集い: 人工知能学会主催) オンライン資料集, pp. 9-15 (2000).
- [Suzuki 00a] Suzuki, R. and Arita, T.: How Learning Can Affect the Course of Evolution in Dynamic Environments, *Proceedings of the Fifth International Symposium on Artificial Life and Robotics*, pp. 260-263 (2000).
- [Suzuki 00b] Suzuki, R. and Arita, T.: Interaction between Evolution and Learning in a Population of Globally or Locally Interacting Agents, *Proceedings of the Seventh International Conference on Neural Information Processing*, pp. 738-743 (2000).
- [鈴木 01a] 鈴木麗聖, 有田隆也: N 人版繰り返し囚人のジレンマゲームにおける空間的局所性の進化, 人工知能学会第 15 回全国大会論文集, 3C2-01 (2001).
- [鈴木 01b] 鈴木麗聖, 有田隆也: N 人版繰り返し囚人のジレンマゲームにおける空間的局所性の影響とその進化, 電子情報通信学会「人工知能と知識処理」研究会

- 報告, Vol. 101, No.66, pp. 39-45 (2001).
- [鈴木 01c] 鈴木麗璽: N 人版繰り返し囚人のジレンマゲームにおける空間的局所性の影響とその動的な進化, 第2回MYCOM (AI若手の集い: 人工知能学会主催) オンライン資料集, pp. 47-52 (2001).
- [Suzuki 01] Suzuki, R. and Arita, T.: Evolutionary Analysis on Spatial Locality in the N-person Iterated Prisoner's Dilemma, *Proceedings of the Inaugural Workshop on Artificial Life*, pp. 105-114 (2001).
- [鈴木 02a] 鈴木麗璽, 有田隆也: 進化と学習の相互作用に対する空間的局所性の影響, 人工知能学会誌, Vol. 17, No. 4, pp. 373-379 (2002).
- [鈴木 02b] 鈴木麗璽: イベント会場における混雑と情報提供に関するマルチエージェントモデル, 第3回MYCOM (AI若手の集い: 人工知能学会主催) オンライン資料集, pp. 33-38 (2002).
- [鈴木 02c] 鈴木麗璽: 行動多様性に対する情報共有の影響とその適応性, 第8回創発システムシンポジウム資料, pp. 107-108 (2002).
- [鈴木 投稿中] 鈴木麗璽, 有田隆也: 行動多様性に対する情報共有の影響とその適応性, 電子情報通信学会論文誌, D-I (投稿中).
- [Suzuki in press] Suzuki, R. and Arita, T.: How learning Can Affect the Course of Evolution in Dynamic environments, *A New Life-style in 21 Century Living with Cognitive and Behavioral Intelligent Artificial Life Robot*, Springer-Verlag (in press).
- [Suzuki accepted a] Suzuki, R. and Arita, T.: Evolutionary Analysis on Spatial Locality in N-Person Iterated Prisoner's Dilemma, *International Journal of Computational Intelligence and Applications* (accepted).
- [Suzuki accepted b] Suzuki, R. and Arita, T.: Effects of Information Sharing on Collective Behaviors in Competitive Populations, *Proceedings of the Eighth International Symposium on Artificial Life and Robotics* (accepted).
- [田村 02] 田村 誠: 社会ゲームと進化—繰り返し「囚人のジレンマ」と適応戦略, 山影 進, 服部正太編, コンピュータの中の人工社会, pp. 91-105 (2002).
- [Todd 96] Todd, P. M.: Sexual Selection and Evolution of Learning, *Adaptive Individuals in Evolving Populations*, Vol. 26, pp. 365-393 (1996).
- [Tuomas 96] Tuomas, W. S., Robert, H. C.: Multiagent Reinforcement Learning in the Iterated Prisoner's Dilemma, *Biosystems* 37, pp. 147-166 (1996).
- [Turney 96] Turney, P., Whitley, D. and Anderson, R. W.: Evolution, Learning, and Instinct: 100 Years of the Baldwin Effect, *Evolutionary Computation*, Vol. 4, No. 3, pp. 4-8 (1996).
- [Waddington 42] Waddington, C. H.: Canalization of Development and the Inheritance of Acquired Characters, *Nature*, Vol. 150, pp. 563-565 (1942).

- [Watson 99a] Watson, R. A. and Pollack, J. B.: How Symbiosis Can Guide Evolution, *Fifth European Conference on Artificial Life*, (1999).
- [Watson 99b] Watson, R. A., Ficici, S. G. and Pollack J. B.: Embodied Evolution: Embodying an Evolutionary Algorithm in a Population of Robots, *Proceedings of the Congress of Evolutionary Computation*, 335-342 (1999).
- [Watts 98] Watts, D and Strogatz, S.: Collective Dynamics of `Small-world' Networks. *Nature*, Vol. 393, pp. 440-442 (1998).
- [Watts 99] Watts, D.: *Small Worlds*, Princeton University Press (1999).
- [山崎 98] 山崎達志, 潮 俊光: Hogg-Huberman 戦略に基づくパケット網ルーチング方式, 信学論 (B) , Vol.J81-B-I, No.3, pp.135-142 (1998).
- [Yamasaki 00] Yamasaki, K. and Sekiguchi, M.: The Clear Explanation of Different Adaptive Behaviors between Darwinian Population and Lamarckian Population in the Changing Environment, *Proceedings of the Fifth International Symposium of Artificial Life and Robotics*, pp. 120-123 (2000).
- [Yao 99] Yao, X. and Darwen, P., An Evolutionary Approach to the N-Player Iterated Prisoner's Dilemma Game, Yao, X. (Ed.), *Evolutionary Computation: Theory and Applications*, 331-357 (1999).

関連発表論文

学術雑誌

- [1] 鈴木麗聖, 有田隆也: 進化と学習の相互作用 - 繰り返し囚人のジレンマゲームにおける Baldwin 効果 -, 人工知能学会誌, Vol. 15, No. 3, pp. 495-502 (2000).
- [2] 鈴木麗聖, 有田隆也: 進化と学習の相互作用に対する空間的局所性の影響, 人工知能学会誌, Vol. 17, No. 4, pp. 373-379 (2002).
- [3] Suzuki, R. and Arita, T.: Evolutionary Analysis on Spatial Locality in N-Person Iterated Prisoner's Dilemma, *International Journal of Computational Intelligence and Applications* (accepted).
- [4] 鈴木麗聖, 有田隆也: 行動多様性に対する情報共有の影響とその適応性, 電子情報通信学会論文誌, D-I (投稿中).

学位論文

鈴木麗聖: 動的環境における進化と学習の相互作用に関する研究, 名古屋大学大学院人間情報学研究科 修士 (学術) 学位論文, 人情修第 326 号 (2000. 3. 27).

口頭発表

a. 査読付き国際会議

- [1] Suzuki, R. and Arita, T.: How Learning Can Affect the Course of Evolution in Dynamic Environments, *Proceedings of the Fifth International Symposium on Artificial Life and Robotics*, pp. 260-263 (2000).
- [2] Arita, T. and Suzuki, R.: Interactions between Learning and Evolution: The Outstanding Strategy Generated by the Baldwin Effect, *Proceedings of Artificial Life VII*, pp.196-205 (2000).
- [3] Suzuki, R. and Arita, T.: Interaction between Evolution and Learning in a Population of Globally or Locally Interacting Agents, *Proceedings of the Seventh International Conference on Neural Information Processing*, pp. 738-743 (2000).
- [4] Suzuki, R. and Arita, T.: Evolutionary Analysis on Spatial Locality in the N-person Iterated Prisoner's Dilemma, *Proceedings of the Inaugural Workshop on Artificial Life*, pp. 105-114 (2001).
- [5] Suzuki, R. and Arita, T.: Effects of Information Sharing on Collective Behaviors in Competitive Populations, *Proceedings of the Eighth International Symposium on Artificial Life and Robotics* (accepted).

b. 国内研究会・シンポジウム・全国大会等

- [1] 鈴木麗璽: 囚人のジレンマゲームにおける Baldwin 効果, **ALIREN** (人工生命研究会) +**ALIST** (在京若手研究者による人工生命自主セミナー) 合同研究会ポジションペーパー (1999).
- [2] 鈴木麗璽, 有田隆也: 囚人のジレンマゲームにおける Baldwin 効果, 人工知能学会第 13 回全国大会論文集, pp. 277-278 (1999).
- [3] 鈴木麗璽, 有田隆也: メタ・パブプロフ: 進化と学習による適応を自動調節する繰り返し型囚人のジレンマ戦略, 情報処理学会研究報告, 99-GI-1, pp. 15-22 (1999).
- [4] 鈴木麗璽, 有田隆也: 学習が進化の過程に与える影響に関するシミュレーション解析, 第 9 回数理生物学シンポジウム予稿集(数理生物学懇談会ニューズレター), p. 9 (1999).
- [5] 鈴木麗璽: 動的環境における進化と学習の相互作用~Baldwin 効果の工学的応用に向けて~, 第 1 回 MYCOM (AI 若手の集い: 人工知能学会主催)オンライン資料集, pp. 9-15 (2000).
- [6] 鈴木麗璽, 有田隆也: N 人版繰り返し囚人のジレンマゲームにおける空間的局所性の進化, 人工知能学会第 15 回全国大会論文集, 3C2-01 (2001).
- [7] 鈴木麗璽, 有田隆也: N 人版繰り返し囚人のジレンマゲームにおける空間的局所性の影響とその進化, 電子情報通信学会「人工知能と知識処理」研究会報告, Vol. 101, No.66, pp. 39-45 (2001).
- [8] 鈴木麗璽: N 人版繰り返し囚人のジレンマゲームにおける空間的局所性の影響とその動的な進化, 第 2 回 MYCOM (AI 若手の集い: 人工知能学会主催)オンライン資料集, pp. 47-52 (2001).
- [9] 鈴木麗璽: イベント会場における混雑と情報提供に関するマルチエージェントモデル, 第 3 回 MYCOM (AI 若手の集い: 人工知能学会主催)オンライン資料集, pp. 33-38 (2002).
- [10] 鈴木麗璽: 行動多様性に対する情報共有の影響とその適応性, 第 8 回創発システムシンポジウム資料, pp. 107-108 (2002).

研究書

Suzuki, R. and Arita, T.: How Learning Can Affect the Course of Evolution in Dynamic Environments, *A New Life-style in 21 Century Living with Cognitive and Behavioral Intelligent Artificial Life Robot*, Springer-Verlag (in press).

調査報告書

複雑系科学の産業応用に関する調査研究報告書 第 2 部 「人工生命手法を用いたマルチエージェント・シミュレーション」担当, 財団法人人工知能研究振興財団 (2002).

