

協調社会の創発を促進する社会メカニズムの構築に向けた進化ゲーム理論的研究

宮地, 浩平

<https://hdl.handle.net/2324/7157378>

出版情報 : Kyushu University, 2023, 博士 (工学), 課程博士
バージョン :
権利関係 :

令和5年度

博士論文

協調社会の創発を促進する社会メカニズムの
構築に向けた進化ゲーム理論的研究

九州大学大学院総合理工学府
環境エネルギー工学専攻
複雑系社会環境科学研究室

宮地 浩平

提出年月日 令和5年7月23日

内容

1.	総論.....	5
1.1.	はじめに.....	5
1.2.	本論の諸研究に至る背景.....	6
1.3.	論文構成.....	10
2.	進化ゲーム理論.....	12
2.1.	ゲーム理論.....	12
2.2.	2 x 2 ゲーム.....	13
2.3.	進化ゲーム.....	14
2.3.1.	進化論.....	14
2.3.2.	ゲーム理論の進化への適用.....	14
2.4.	ネットワーク互惠.....	15
2.4.1.	複雑ネットワーク.....	15
2.5.	ネットワークパラメーター.....	16
2.5.1.	ノード数.....	16
2.5.2.	次数（リンク数）.....	16
2.5.3.	平均ノード間距離.....	18
2.5.4.	クラスター係数.....	18
2.6.	戦略適応方法.....	18
2.6.1.	イミテーションマックス（IM）.....	18
2.6.2.	ペアワイズ（PW）.....	19
2.6.3.	ルーレット選択（RS）.....	19
2.7.	ネットワークの適応.....	19
2.8.	Ultimatum Game の基礎.....	20
2.8.1.	Ultimatum Game におけるナッシュ均衡とパレート最適.....	20
2.8.2.	アダプティブダイナミクス.....	21
2.8.3.	Ultimatum Game でのアダプティブダイナミクス.....	22
3.	Ultimatum Game におけるネットワークと戦略適応の共進化モデルに基づく社会的協調の自己組織化.....	25
3.1.	研究背景と着眼点.....	25

3.2. 研究概要.....	25
3.3. モデル.....	26
3.3.1. 拡張 Ultimatum Game	26
3.3.2. 離散戦略.....	27
3.3.3. ネットワーク適応における切断プロトコル.....	27
3.3.4. 各切断プロトコルの詳細.....	28
3.3.4.1. 確定的切断プロトコル.....	28
3.3.4.2. 確率的切断プロトコル.....	29
3.3.5. シミュレーション設定.....	30
3.4. 結果と考察.....	31
3.4.1. 次数分布.....	31
3.4.2. 戦略値進化の考察.....	33
3.4.2.1. $b = 1.0$ の場合.....	33
3.4.2.1.1. 不合理切断の生起回数.....	35
3.4.2.2. $b = 2.0$ の場合.....	37
3.4.3. ネットワークパラメーター.....	39
3.4.4. 離散戦略.....	42
3.4.4.1. $b = 1.0$ の場合.....	42
3.4.4.1.1. 平衡状態.....	44
3.4.4.2. $b = 2.0$ の場合.....	45
3.4.5. 離散戦略での均衡時におけるネットワークパラメーター.....	46
3.5. 結論.....	49
4. 記憶に基づいた確率的意思決定が協調行動に与える影響.....	50
4.1. 研究背景と着眼点.....	50
4.2. 研究概要.....	50
4.3. モデル.....	50
4.3.1. 2×2 ゲーム.....	50
4.3.2. シミュレーション手順.....	51
4.4. 結果と考察.....	52
4.5. 結論.....	60

5.	混合戦略モデルにおける戦略更新方法に内包される確率的揺らぎと協調創発の関係	62
5.1.	研究背景と着眼点.....	62
5.2.	研究概要.....	62
5.3.	モデル.....	62
5.3.1.	2×2 ゲーム.....	62
5.3.2.	ネットワーク.....	63
5.3.3.	戦略更新と行動更新.....	63
5.3.4.	シミュレーション設定.....	64
5.4.	結果と考察.....	64
5.4.1.	(補足) 均衡協調率と平均利得の関係.....	72
5.4.2.	(補足) クラスタ中に生じる特殊な状態.....	73
5.5.	結論.....	74
6.	中間(日和見)戦略の定義による囚人のジレンマ解消メカニズムの解明.....	76
6.1.	研究背景と着眼点.....	76
6.2.	研究概要.....	76
6.3.	モデル.....	76
6.3.1.	2×2 ゲーム.....	76
6.3.2.	シミュレーション手順.....	77
6.4.	結果と考察.....	77
6.5.	結論.....	84
7.	ネットワーク上の同類的な繋がり の促進が囚人のジレンマゲームにおける協調進化に与える影響.....	86
7.1.	研究背景と着眼点.....	86
7.2.	研究概要.....	86
7.3.	モデル.....	86
7.3.1.	2×2 ゲーム.....	86
7.3.2.	共進化モデルにおけるシミュレーション.....	87
7.3.3.	ゲーム.....	87
7.3.4.	戦略適応.....	88

7.3.5. ネットワーク適応—同類的接続	88
7.4. 結果と考察	88
7.5. 結論	97
8. 総括結論	98
8.1. 結論	98
8.2. 今後の展望	100
参考文献	103
謝辞 107	

1. 総論

1.1. はじめに

社会では、個人の利益を最大化する行動と社会の利益を最大化する行動に乖離がある場合があり、こういった状況を社会ジレンマ構造と呼ぶ。意思決定プロセスを静学的アプローチによって論理的に記述する理論としてゲーム理論がある。ゲーム理論は応用数学の一種として様々な分野の学問に応用されており、主に経済学で用いられている他、社会行動学、心理学、生物学、物理学、政治学、軍事学、環境工学などにも用いられている。ジレンマ環境下において、古典ゲーム理論では個人の利益を最大化するために最も合理的とされている行動が導き出されている。一方で、現実社会では必ずしも古典ゲーム理論で予測されているような個人の利益のみを最大化するような行動はとられず、互いに協調しあって社会全体にとって最適な行動がとられている場面が散見されるが、現実社会においてどのようにして協調的行動が創発しているのかという点に関してはいまだに謎に包まれている。

生物学におけるゲーム理論の発展理論として非常に有名なのが 1982 年に Maynard Smith[1] によって提唱された進化ゲーム理論である。生物の進化の概念は、Darwin による進化論の登場以降、遺伝、変異、淘汰の三つの要素からなると理解されている。種は自分のコピーを後生に遺伝することで基本的に種の存続をはかるが、そのコピーは精密なものではないため、変異によってその種から派生する種が登場する。それら種の存続能力にはそれぞれ差があり、その差による競争がまさに淘汰を生む。この進化の過程を力学系ダイナミクスとして捉えた研究は数多く存在するが、そしてその中の一つが進化ゲームである。先述したとおり、古典ゲーム理論それ自体に基づく予測と現実社会における現象は一致しないことがあるが、進化ゲーム理論では、ダーウィンの自然選択説に基づくいくつかの仮定を導入することで、進化のダイナミクスをうまく表現している。進化ゲーム理論の利点は、古典ゲーム理論で仮定されていた意思決定主体の「合理性」を仮定せずとも、相互依存状況を再現できることにある。

進化ゲーム理論の中でも近年注目されているものの 1 つとして、囚人のジレンマゲーム (PDG ; Prisoner's Dilemma Game) における協調行動の創発メカニズムの研究が挙げられる。Nowak [2]は 2×2 ゲームを基本とした一連の理論研究により、協調創発は直接互惠、間接互惠、血縁選択、群淘汰といった 5 つの基本的互惠機構に類別出来ることを示した。

本論では、主に進化ゲーム理論モデルにおけるネットワーク互惠の影響について考察していく。人と人とのつながりが協調創発に寄与しているとの視点に立つネットワーク互惠

では人間社会システムだけでなく生物システムにおける互惠関係の進化を説明する有力なモデルであるとの認識のもと、非常に多くの研究が統計物理学、数理生物学、情報学分野で行われてきている。ネットワーク互惠を加味した進化ゲーム（ネットワークゲーム）では、エージェントは潜在的に協調（C）か裏切り（D）かの2戦略いずれか、ネットワークで接続された隣人全てとゲーム対戦して、毎回ゲーム後に隣人の中からあるプロトコルに従ってC戦略もしくはD戦略をコピーすることで戦略適応を行う。エージェント同士のつながりを規定する基盤ネットワークや戦略適応プロトコル、それらの組み合わせは、ほぼ無数に仮定できるため、上記したようにこれまで数多くのモデル研究が行われてきたわけだが、近年これらの影響を俯瞰的に考察し、ネットワークの互惠ダイナミクス上の本質的効果を解明しようとの試みが報告されてきている[3][4][5]。

ネットワーク互惠と協調行動創発の関係を考察する研究は静的ネットワークを仮定して行うことが一般的である。一方で、人と人とのつながりを規定するネットワークは常に固定させているわけではなく、動的に進化していくものであるという考え方がある[6][7]。あるルールに従ってネットワークそのものも動的に進化していく過程を導入することで系での協調の進化が促進されることがわかっている。本論では、動的なネットワークを仮定した共進化モデルを用い、協調行動の創発メカニズムについて詳しく考察していく。

研究手法としてはC++を用いた数理シミュレーションモデルを構築し、進化ゲーム理論をベースにさまざまな環境変数を導入することで、どのような要素が協調行動の創発に寄与し得るのかを実験的に検証する。進化ゲーム理論をもととしたこの数理モデルは人と人との相互作用をゲーム理論の概念を適用して極限まで単純化して表現したものであるため、解析対象として特定の現象に対する限定性を持たない。つまり、実社会におけるあらゆる現象に対して発展・応用の可能性を秘めているといえる。本論では現実に存在する人々の相互作用を進化ゲーム理論にて表現する中で、進化ゲーム理論に基づく数理モデルそのものの挙動に関する数学的な分析に加えて、もたらされた結果が実社会において何を示唆しているのとらえることができるのかについても考察していく。

1.2. 本論の諸研究に至る背景

自己利得最大化のために利己的行動を採ると社会全体効用が低下する社会ジレンマは環境問題をはじめ様々な場面で観察される。人間やその他の生物種では、各自が他利的行動を採り協調を自己組織化することで、社会ジレンマを克服している。こういった現象を解明する

ため、進化ゲーム理論研究で用いられる標準的テンプレートとしては先に述べた 2×2 (2人2戦略) ゲームが有名である。 2×2 (2人2戦略) ゲームでは、自身の戦略的意思決定と相手の選択、それによってもたらされる自身および社会全体の利益の相互関係を表現するモデルである。 2×2 (2人2戦略) ゲームとは別のゲーム理論研究の枠組みとして **Ultimatum Game** (最後通牒ゲーム) がある。**Ultimatum Game** では、経済的取引を行う場面を仮定したゲームモデルである。有限の資産を提案者の提案した金額割合で分け合い、受諾者が受け入れれば提示金額で取引が成立、拒否すれば取引は不成立となり、提案者、受諾者ともに利益は得られない。この枠組みでは、各人のもつ公平性の概念、利己的な動機および合理性を表現・研究するモデルとなっており、Güth ら[8]の研究以降、**Ultimatum Game** は社会的疑問や人間の振る舞いを研究するためのツールの一つとして用いられてきた(詳細は2章を参照のこと)[9][10][11][12].

本論では、本論では、 2×2 (2人2戦略) ゲームに加えて **Ultimatum Game** を研究題材として取り上げる。これは、①人間社会における非対称な相互関係を記述でき、②人間意志決定機構を模擬するに適切である連続値戦略空間[13][14]でゲームが定義できるとの事由による。加えて、進化ゲームとしては異なる枠組みである 2×2 (2人2戦略) ゲームを用いて研究されてきた互惠機構が **Ultimatum Game** に対しても有効であるのかを検証することで、両進化ゲーム理論研究の成果を統合し新たな知見を得たいという意図がある。**Ultimatum Game** の枠組みを用いることで人間が持ちうる公平性の創発事由を示唆できることから、実験的なアプローチと論理的アプローチ双方での研究が盛んにおこなわれている[15][16][17][18][19][20][21]. **Ultimatum Game** において非常に興味深い点は、全員の匿名性が確保されている所謂 **Well-mixed** の状態においては、理論的にはエージェントはどんな小さなオファーでも受け入れ、相手には最小のオファーを行うというのが合理的な選択として導き出されている一方で、実証実験や過去の研究に基づく、実際には約 50%の割合での資本分割をオファーし、自身には全くの利益が無いにもかかわらず、30%以下の割合での分割オファーは拒否するという非合理的な選択を人間がとる、ということがわかっている点である。過去の研究では、**Reputation** (周囲からの評判)、共感、コストと罰則といった社会科学的観点に加えて、ネットワーク構造による匿名性の減少の観点からの説明も試みられている[22][23][24][25][26][27][28][29][30][31]. 本論では、**Ultimatum Game** における一般的な戦略進化機構に加え、固定グラフ上のネットワークゲームに、所謂、 2×2 (2人2戦略) ゲームで研究が盛んにおこなわれてきている **Game Exit Option**[21]を付加したネットワーク進化の機構(戦

略とネットワークの共進化) を考慮することで公平性の創発原理の説明を試みる。ネットワークが動的に進化する機構によって公平性の創発を説明する既往研究はいくつかあるが[32][33][34][35], 依然として動的なネットワーク進化の機構を定義づけるネットワーク進化ルールの規定と公平性の創発における関係性にはさらなる研究の余地がある。本論では, 動的ネットワークの進化ルールを複数仮定し, 2×2 (2人2戦略) ゲームで得られている知見を組み合わせ, 公平性の創発につながるネットワーク進化機構を探求する。

2×2 (2人2戦略) ゲームによって定義される社会数理ジレンマ構造は複数ある(詳細は2章参照)。通常, 2×2 (2人2戦略) ゲームによってジレンマの解消を試みるアプローチとしては, 一つのジレンマクラスを選定し, ジレンマの解消につながると思われるメカニズムを導入して進化の過程を分析していくという手法をとる。ジレンマ解消のためのアプローチとしては, 先に述べたネットワーク構造の違いによる影響や[36][37], 戦略適応方法による違い[38][38][39][40][41][42][43][44][45][46][47][48][49][50][51]を考察することが一般的である。こういった研究の際には囚人のジレンマゲームが研究対象のジレンマクラスとして選ばれることが多い。その理由としては, 囚人のジレンマという数理ジレンマは現実社会におけるビジネス競争や国際政治環境といったものに相似性を見出しやすいからである。一方で, もう一つの有名な数理ジレンマクラスを体現するものとして Snow Drift (もしくは Chicken ; チキン) Game がある[52]。チキンゲームは, 軍事的対立のエスカレートによる状況の悪化や, 市場における価格競争といった状況を模擬したジレンマクラスであるといえることができるだろう。チキンゲームにおいては, 双方がむきになって強硬(裏切)戦略をとり続けた場合, 双方にとって最も悪い状況となる。どちらか一方が譲歩した場合, 譲歩した側(譲歩戦略(協調戦略)側)の得られる利益は限られるものの, 譲歩を勝ち取った側(強硬戦略(裏切戦略)側)は極めて大きな利益を得られる。こういった状況下では, 囚人のジレンマゲームにおいては最適な状況であった互いに協調しあう状況よりも, 交互に協調戦略と裏切戦略を取り合う方が社会全体として見たときに結果的に大きな利得を得られることがある。これを囚人のジレンマにおける R 互惠[53][54]に対して, チキンゲームにおける ST 互惠と呼ぶこととする。このように数理ジレンマクラスによって社会にとって好ましい社会の帰結は異なってくる。異なるジレンマクラスにおける普遍的なジレンマ解消メカニズムという視点での研究はあまり行われてきていないことから[55], 本論では, 5つの互惠機構の一つである直接互惠[56]をネットワークゲームモデルに組み込むことで, 異なるジレンマクラスにおいてそれぞれ好ましい社会的帰結を可塑的にもたらせるような機構の解明を試みる。実社会において観察されるジレ

ンマクラスは先の例に述べたように一つとは限らない。複数のジレンマクラスに対する耐性を持ったユニバーサルな機構を探求することには複雑社会における協調創発メカニズム解明の一手として意義があると考えられる。

2×2 (2人2戦略) ゲームにおいて、戦略的意思決定に内包されるランダム過程 (確率的要素) が協調の創発に寄与しているとの報告がある[57]。既存の枠組みのように一意に協調または裏切り戦略をとり続けるのではなく、自身の保有する実数戦略値に基づき確率的に協調または裏切り戦略を決定する混合戦略モデル[58]では、系の帰結は一般的な離散戦略モデルと大きく異なるものであることがわかっている。確率的要素が協調の創発に寄与するという点は、進化ゲームモデルで仮定する基盤ネットワークや戦略更新ルールに内包される確率的特性に関する研究からも関連性があることが導き出されている[59][60][61][62][63][64]。本論では、混合戦略モデルによる戦略更新過程の確率的な振動・揺らぎといった要素が数理モデル中でどのように作用し、協調の創発に影響しているかを詳細に考察する。また、従来は離散戦略モデルと混合戦略モデルは完全に別の戦略定義体系として捉えられていたが、本論では離散戦略モデルと混合戦略モデルを連続的に定義できる戦略解像度という概念を新たに導入し、従来の離散戦略と完全な混合戦略の中間的混合戦略といえる戦略の定義を可能にすることによって、混合戦略モデルによるエージェントの確率的振る舞いが協調創発に寄与するメカニズムの本質的な理解を試みる。

Ultimatum Game で仮定している動的ネットワークに関しては、2×2 (2人2戦略) ゲームにおける研究でも盛んに研究が行われている。動的ネットワークの考え方は複数あり、代表的な考え方としてネットワークをあるルールに従って戦略とともに進化させていく共進化モデルがある[65][66][67][68][69][70][71]。他には、格子ネットワーク上にあえて空隙を設けておき、エージェントがルールに従って空隙格子状を移動していく Mobility モデル[72][73][74][75][76]や、自身とつながりを持つ隣人エージェントとの接続数を増減させることでネットワーク構造自体が成長していく成長ネットワークモデル[77][78][79]、複数のネットワークタイプを系に仮定する Interdependent Network Model[80][81][82]等がある。こういった既往研究は基本的に離散戦略モデルを用いて行われてきた。本論では、混合戦略モデルにおける共進化動的ネットワークを仮定することで、協調創発メカニズムに関して新たなアプローチによる解明を試みる。混合戦略モデルにおいては、実数戦略値である特性を生かして従来の離散戦略モデルでは定義できないような柔軟なネットワーク進化ルールを規定できるため、協調創発メカニズムに関する新たな知見の獲得が期待できる。

本論では、上記で解説したように、既往研究のさらなる発展的および新たな視座を以って包括的に進化ゲーム理論モデルを分析することで、数理モデルそのものに対する理解をより深いものとし、新たな協調進化の機構を解明することによって、進化ゲーム理論研究分野の発展に貢献することを目的とする。

1.3. 論文構成

第 2 章では本論の基礎論理フレームワークとしてゲーム理論の基礎、進化ゲーム理論研究に用いられる 2×2 (2 人 2 戦略) ゲームおよび *Ultimatum Game* (最後通牒ゲーム) と戦略適応に関する枠組みと、本論の根幹となるネットワーク互惠と複雑ネットワークに関して解説する。第 3 章では最後通牒ゲーム (*Ultimatum Game*) において共進化ネットワークを仮定し、各エージェントの保有する戦略値に基づきいくつかのネットワーク切断プロトコルを仮定する。各ネットワーク切断プロトコルにおけるネットワークおよび戦略値の進化の帰結を分析する。また、資本分割オファーに対する増幅係数を仮定し、こういった要素が進化の帰結及び協調創発のメカニズムに関してどのように作用するのか分析を行う。第 4 章では、ネットワーク互惠に加えて、進化ゲームにおける記憶長の効果 (直接互惠) を加味する。実数値で定義された記憶長を加味することによって、囚人のジレンマ域では R 互惠、チキンゲーム域では ST 互惠といった、各ジレンマクラスにおいて社会全体として最も好ましい進化の帰結を達成できることを説明する。第 5 章では、混合戦略モデルにおける戦略更新ルール (シンクロ更新もしくはアシンクロ更新) とネットワークの次数の各組み合わせを考察する中で、確率共鳴によって高次数・強ジレンマ域においても協調が創発するメカニズムについて詳述する。第 6 章では、混合戦略ゲームにおける戦略解像度を定義することで、離散戦略と完全混合戦略を連続的に定義し、混合戦略に内報される確率的要素が協調の創発に与えている影響に関して考察する。第 7 章では、混合戦略モデルを用い、自身の戦略値をエージェントの特性 (個性) と仮定し、同類的な繋がり (自身と類似する戦略をもつ者同士との繋がり) を促進する動的ネットワーク進化ルールによる協調の進化のダイナミクスについて述べる。3 章から 7 章までの各研究の関連を表 1.3.1 に示す。最後に 8 章にてまとめと展望を述べる。

	3章	4章	5章	6章	7章
2x2ゲーム		○	○	○	○
最後通牒ゲーム	○				
ネットワーク互恵	○	○	○	○	○
直接互恵		○			
動的ネットワーク	○				○
混合戦略		○	○	○	○

表 1.3.1 本論における各章と研究着眼点マトリックス

2. 進化ゲーム理論

2.1. ゲーム理論

ゲーム理論では複数の意思決定主体が存在する中で、連携的ではなく個人的な目的で合理的に意思を決定する非協力ゲームを考える。本論ではこの個体をエージェント i 、エージェントの取る行動の根拠となるものを戦略 S_i と呼ぶ。エージェントはゲームにおいてある戦略を取ることで何かしらの結果を得るが、この結果を利得と呼ぶ。ゲーム理論における利得はエージェント同士の相互作用によってそれぞれにもたらされる実質的利益、損得、感情的な帰結を数字で表したものである。非協調ゲームにおける合理的な行動とは、この利得を最大にするように戦略を取ることに相当する[1]。

戦略には大きくわけて離散戦略と混合戦略がある。前者は行動の選択肢一つ一つを意味し、混合戦略は、その選択肢をどう取るかを意味する。じゃんけんを例にとれば、グー、チョキ、パーの夫々が純粋戦略であり、それを全て $1/3$ の確率で出すことが混合戦略に相当する。

現実的には対戦するエージェントの数は多数いることも考えられるが、非協力ゲームの場合対戦相手は一度に一人であることが多い。ここでエージェント 1（自分）の戦略が $S^1_1, S^1_2, \dots, S^1_m$ の m 個、エージェント 2（相手）の戦略が $S^2_1, S^2_2, \dots, S^2_n$ の n 個の場合を $m \times n$ ゲームという。この場合お互いの取る戦略の組み合わせによって決まる利得を、表 2.1 のように $m \times n$ 行列に表すことが出来る。

1 \ 2	S^2_1	S^2_2	...	S^2_n
S^1_1	$u^{1_{11}}, u^{2_{11}}$	$u^{1_{12}}, u^{2_{12}}$...	$u^{1_{1n}}, u^{2_{1n}}$
S^1_2	$u^{1_{21}}, u^{2_{21}}$	$u^{1_{22}}, u^{2_{22}}$...	$u^{1_{2n}}, u^{2_{2n}}$
:	:	:	...	:
S^1_m	$u^{1_{m1}}, u^{2_{m1}}$	$u^{1_{m2}}, u^{2_{m2}}$...	$u^{1_{mn}}, u^{2_{mn}}$

表 2.1 $m \times n$ ゲームの利得表

合理的に自分の戦略を決めようとするとき、相手の戦略によって自分の最適戦略、つまり利得を最大にする戦略は異なるし、当然相手もその状況に置かれている。ここで、相手が「ある戦略」をとっているときの自分の最適戦略において、相手にとっての最適戦略が「ある戦

略」であるとき、これをナッシュ均衡と言う。このナッシュ均衡はどのようなゲームに置いても必ず存在するが、必ずしも一つとは限らない。表 2-2 にある 3×3 ゲームにおけるナッシュ均衡の例を示した。

2 1	S^2_1	S^2_2	S^2_3
S^1_1	4,4	3,1	2,3
S^1_2	1, 3	2,2	4,2
S^1_3	3,2	2,4	5,5

表 2.2 利得行列上でのナッシュ均衡探索

太字で示した要素がそれぞれの相手の戦略に対しての最大利得を示している。この例を詳しく見てみる。例えば仮に相手が S^2_1 の戦略をとると仮定する。その場合自分は S^1_1 をとることが最適である。このことは相手にも言えるため、 S^1_1 , S^2_1 の組み合わせはナッシュ均衡である。次に相手が S^2_2 の戦略をとる場合を考える。この場合自分は S^1_1 を表とることが最適であるが、相手からするとこの S^1_1 を取る場合 S^2_2 をとることは適当ではない。このように考えると、ナッシュ均衡であるためには同一ブロックで両方とも太字となっている場合であることがわかる。よってこの場合は S^1_3 , S^2_3 の組み合わせもナッシュ均衡である。

表 2.2 を更に詳しく見てみる。両者にとって理想的な帰結は、お互いが 5 の利得を上げることが出来る S^1_3 , S^2_3 の組み合わせである。一般に、全エージェントにとって状態 A の方が状態 B より望ましいとき、A は B よりパレート優位であるという。更に A が全ての状態において最良の状況であるとき A はパレート最適であるといい、この例ではまさに S^1_3 , S^2_3 の組み合わせがパレート最適である。しかしこの例ではこの組み合わせ以外にもナッシュ均衡が存在している。このように、パレート最適以外にも均衡を持つ状況を、非協調ゲームではジレンマという。

2.2. 2×2 ゲーム

$m \times n$ ゲームにおいて $m=2$, $n=2$ であり、さらに両者条件が同じである場合を 2×2 ゲームという。ゲーム理論では伝統的にエージェントの取り得る戦略（行動）の種類を協調戦略と裏切り戦略の二種類に限定する。以下本論ではエージェントの戦略は協調、または裏切りの二種類

であるとする。この二種類の戦略同士の相互作用を規定した利得表を表 2.3 に示す。なお表中の戦略 C, D は夫々 Cooperation, Defection の頭文字を表す。

	C	D
C	R,R	S,T
D	T,S	P,P

表 2.3 2×2 ゲームの利得表

ここで R は Reward, S は Sucker, T は Temptation, P は Punishment の頭文字を夫々表している。両者が同じ条件下にいる場合は、行列が対称になるために、表 2.4 のように簡単に表すことができる。

	C	D
C	R	S
D	T	P

表 2-4 簡易表現した 2×2 ゲームの利得表

以後本論ではこのような表現をした場合は対称ゲームを示しているものとする。

2.3. 進化ゲーム

2.3.1. 進化論

生物の進化の概念は、Darwin による進化論の登場以降、遺伝、変異、淘汰の三つの要素からなると理解されている。説明を補足すると、種は自分のコピーを後生に遺伝することで基本的に種の存続をはかるが、そのコピーは精密なものではないため、変異によってその種から派生する種が登場する。それら種の存続能力にはそれぞれ差があり、その差による競争がまさに淘汰を生む、との理解である。この進化の過程は状態の時間発展に他ならず、これを力学系ダイナミクスとして捉えた研究は数多く存在する。そしてその中の一つが進化ゲームである。

2.3.2. ゲーム理論の進化への適用

ゲーム理論においてゲームを構成しているのはエージェント、戦略、利得のみである。この

うち、エージェントは種を構成する個体に対応する。残るは戦略と利得だが、これらは一体何を意味するのか。進化論における三要素のうちゲーム的状况で表現できるのは淘汰であると考えるのが自然である。とすれば種間競争による適応度の差が戦略間のゲームにおける利得によって決められると考えられる。ここで進化に関する議論においてしばしば適応度の定義としては、少なくとも共通していえるのが「殖えやすいものは殖える」という概念である。よって、取得利得が相対的な適応度の大小に相当する。またゲーム理論では、ある状況における合理的行動を目的としたが、進化における種はエージェントにとって不変であることから、戦略はエージェントそれぞれが生まれながらにして所有しているもの、例えば種や遺伝子に相当するとし、合理性は考えない[1][83].

2.4. ネットワーク互惠

2.4.1. 複雑ネットワーク

ネットワークという言葉は、ノードとそれを繋ぐリンクが存在する所謂網目構造を持つものの全般に使われている。例えばこれらにはバスや列車のような交通網、物体の分子構造など空間構造を持つものから、ウェブサイトや人間関係など、相互関係性を表すものにも用いられる。こういった相互関係性をネットワークと言うならば、ほぼ全てのものにはネットワーク構造が存在することになる[84]。近年では SNS 等のソーシャルメディアでの情報拡散メカニズム等の社会的現象、最適なエネルギーグリッド構成等に関してもネットワーク理論を適用した手法による分析が行われている。

このネットワークという学問のベースとなるグラフ理論と呼ばれる分野の研究の歴史は古く 18 世紀まで遡るが、近年再び盛んに研究が行われるようになってきている。火付け役となったのは Erdos と Renyi [85] によるランダムグラフと Watts と Strogatz [86] によるスモールワールドの考案などがある。

複雑ネットワークは学問そのものとしても非常に面白い上に、社会現象や本研究で扱う進化生物学などにも応用できるため、ネットワークを詳しく説明するだけで膨大な量になる。しかし本研究はネットワークそのものが研究対象ではないため、ここでは中でも重要と思われるネットワークパラメーターとグラフを説明することに留める。

2.5. ネットワークパラメーター

2.5.1. ノード数

ノード数はまさに読んで字のごとくネットワークを構成する全体のノードの数を意味し、 N で表す。ネットワークの例を図 2.3 に示したがこの例では $N = 10$ となる。

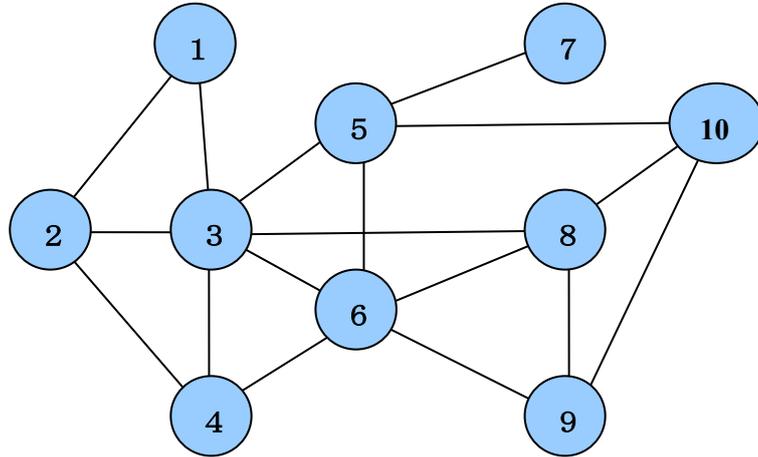


図 2.3 ネットワークの例

2.5.2. 次数（リンク数）

次数とはリンク数と同じ意味で、直接繋がっているノードの数であり、 k で表す。図 2.3 のノード 1 では $k=2$ である。

また全体が平均的に幾つのノードと繋がっているかを表したものを平均次数といい

$$\langle k \rangle = \frac{\sum_{i=1}^N k_i}{N} \quad (2.5.1)$$

で表される。図 2.3 では $\langle k \rangle = 3.4$ となる。

次に次数の確率分布 $p(k) = \{p(1), p(2), \dots\}$ を考える。 K を最大次数とすると当然

$$\sum_{k=1}^K p(k) = 1 \quad (2.5.2)$$

が成立している。図 2.3 の場合の次数分布を図 2.4 に示した。

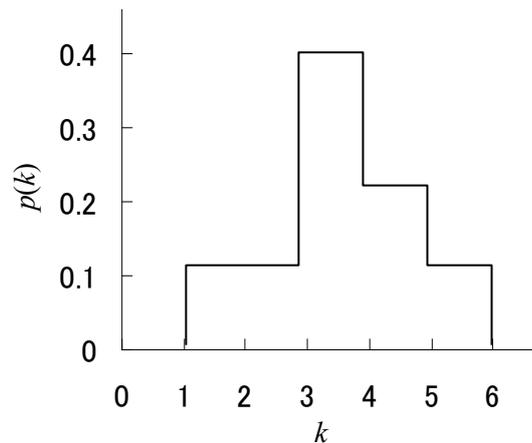


図 2.4 図 2.3 のネットワークの次数分布

更に次数に関するパラメーターで 2002 年 Newman[87]の報告により注目されるようになったものに次数相関がある。これは繋がっているノード同士の次数の傾向を表すパラメーター r で表される。まずリンクで繋がっているノード同士が、そのリンク以外にあと何本リンクを所有しているかを考える。これを残余次数と呼ぶことにする。この残余次数分布はノードが所有している次数より 1 少ないため強度分布が $(k+1)p(k+1)$ と等価である。よってノーマライズした残余次数強度分布 $q(k)$ は $p(k)$ を用いて

$$q(k) = \frac{(k+1)p(k)}{\sum_j jp(j)} \quad (2.5.3)$$

と表される。次にリンクで繋がっているノードの残余次数の接続確率分布を e_{jk} で定義する。ここで j と k には対称性が保たれるため $e_{jk}=e_{kj}$ が成り立つ。この e_{jk} は次の二式を満たす。

$$\sum_j \sum_k e_{jk} = 1 \quad (2.5.4)$$

$$\sum_j e_{jk} = q(k) \quad (2.5.5)$$

ここで両者の残余次数に相関がなければ $e_{jk}=q(j)q(k)$ となるが、相関があればそうはならない。よってこの相関を $\sum_j \sum_k jk(e_{jk} - q(j)q(k))$ とし、これをノーマライズした

$$r = \frac{1}{\sigma_q^2} \sum_j \sum_k jk(e_{jk} - q(j)q(k)) \quad (2.5.6)$$

を次数相関と定義する。ここで

$$\sigma^2 = \sum_k k^2 q(k) - [\sum_k k q(k)]^2 \quad (2.5.7)$$

である。さらに次数相関の式に対してネットワーク情報が与えられているときに r を評価する場合には

$$r = \frac{M^{-1} \sum_i j_i k_i - [M^{-1} \sum_i \frac{1}{2}(j_i + k_i)]^2}{M^{-1} \sum_i \frac{1}{2}(j_i + k_i) - [M^{-1} \sum_i \frac{1}{2}(j_i + k_i)]^2} \quad (2.5.8)$$

を用いることも出来る．ここで $i=1, \dots, M$ のリンク番号を表し, j_i, k_i はそれぞれリンクの両端のノードの次数を表す．この r は $-1 \leq r \leq 1$ の範囲の値をとり, r が正であればノードは近い次数同士で繋がっており, 0 であれば無相関, 負であれば次数の離れたノードと繋がる傾向にあるネットワークであることを示す．現実のネットワークの多くは負の相関を持っているが, 人間社会のネットワークは正の次数相関を持つことがわかっている．

2.5.3. 平均ノード間距離

ノード間距離とはネットワーク上の任意の2点が最低何リンクを介して到達するかを表す．図 3.1 のノード 1 と 10 のノード間距離は 3 となる．これを全ノードの組み合わせで平均したものを平均ノード間距離といい, L で表す．図 2.3 では $L \cong 1.8$ となっている．

2.5.4. クラスタ係数

人間の友達関係において, 自分の友達はお互い友達同士であることが多い．ネットワークではこのような関係をクラスタという．例えば図 2.3 においてノード 1, 2, 3 はクラスタを形成している．そして友達同士が友達である割合をそのノードのクラスタ係数といい, C_i で表す．図 2.3 において $C_1=1, C_2=2/3$ である．またこれらを全ノードで平均したものをネットワークのクラスタ係数とし C で表す．図 2-3 では $C = 7/15$ である．

2.6. 戦略適応方法

進化ゲーム理論においては自身の行動を規定する戦略パラメーター (C もしくは D) が時間の経過とともに進化していく戦略適応方法を規定する必要がある．戦略適応方法の違いによって進化の帰結に影響を及ぼすことがわかっている．ここでは本論に用いられる代表的な戦略適応方法を述べる[89].

2.6.1. イミテーションマックス (IM)

自分と自分の隣人のうち最大利得の戦略をコピーする． N_i を自分の隣人とし, 合計利得を Π としたとき, 戦略変化は次の式で表すことができる．

$$S_i = \begin{cases} S_i & \text{if } \Pi_i > \max\{\Pi \in N_i\} \\ S_j & \text{if } \Pi_j = \max\{\Pi \in N_i\} \end{cases} \quad (2.6.1)$$

2.6.2. ペアワイズ (PW)

自分の隣人のうち一人をランダムに選択し、確率的にその相手の戦略をコピーする。その確率の決め方は大きく二通りあり、一つは利得差をフェルミ関数で評価して決めるもので

$$W_{S_i \leftarrow S_j} = \frac{1}{1 + \exp[(\Pi_i - \Pi_j)/\tau]} \frac{1}{1 + \exp[(\Pi_i - \Pi_j)/\tau]} \quad (2.6.2)$$

の確率で相手の戦略をコピーする (以下 F-PW)。なお式中の τ は物理学でいう温度に相当するランダム性を表すパラメーターである。もう一つは利得差を線形的に評価して決めるもので

$$W_{S_i \leftarrow S_j} = \frac{\Pi_j - \Pi_i}{\max(k_x, k_y)[\max(R, T, S, P) - \min(R, T, S, P)]} \quad (2.6.3)$$

の確率で相手の戦略をコピーする (以下 L-PW)。分母は確率が 0~1 の範囲に収まるよう規格化するためのもので、両者の次数のうち大きいものに利得行列の要素で最大のもものと最小のものとの差を掛けたものである。なお、本論で検討されている **Ultimatum Game** では最大利得は $p_i=0$, $p_j=1$ のときの 2 である。

2.6.3. ルーレット選択 (RS)

ルーレット選択では自分若しくは自分の隣人のうち一人を獲得利得に応じて確率的に選択しコピーする方法。ただしこのとき確率が負値にならないように、計算では自分と隣人のうち最低利得者の利得を全エージェントの利得から引いた利得を用い

$$W_{S_j \leftarrow S_i} = \frac{\Pi_j - \min_{k \in N_i} [\Pi_k]}{\sum_{j \in N_i} (\Pi_j - \min_{k \in N_i} [\Pi_k])} \quad (2.6.4)$$

と表される確率でコピーする。

2.7. ネットワークの適応

戦略の適応だけでなく、エージェントのつながりを規定するネットワークそのものも動的に進化していくという共進化という考え方がある。ネットワーク適応によって進化ゲームにおける協調の創発に寄与するといった研究成果がある (例えば[65][66][67][68][69])。ネットワーク適応の方法は研究によってさまざまである。例えば、フォーカルの隣人の中で最も利

得が低いエージェントのネットワークを切断する，隣人の中で最も協調率が低いエージェントを切断する，などといった切断方法がある．切断を行ったエージェントは，新たな相手とランダムに接続する場合もあれば，フォーカル近傍のエージェントと接続する場合もあり，この処理も多種多様である． **Game Exit Option** として共進化メカニズムを研究した代表的な研究成果として[21]がある．

ネットワークの切断・接続プロトコルの違いは最終的な均衡点やネットワークパラメータに大きく関係してくる．また，初期に生成する基盤のネットワークが均衡時のネットワークパラメータに影響を及ぼすとの研究もある（例えば[90]）．

一般的に戦略適応とネットワーク適応の共進化モデルを定義する際，パラメータ w を導入し， w の確率でネットワーク適応， $1-w$ の確率で戦略適応を行うと定義することで， w が大きくなるほどネットワーク適応速度が上がる状態を表現する[34]．すなわち， $w=0$ の場合はネットワーク適応が生起せず，戦略適応のみが生起する既往の固定ネットワークモデルを表す．

2.8. Ultimatum Game の基礎

2.8.1 Ultimatum Game におけるナッシュ均衡とパレート最適

Ultimatum Game は， 2×2 ゲームと同様，ゲーム理論での議論において頻出するモデルである．Ultimatum Game では，エージェント i は $[0,1]$ の実数で定義された二次元戦略値 $S_i=(p_i, q_i)$ を有する．1回のゲーム対戦は各エージェントが **Proposer** と **Responder** になる2局より構成される．エージェント i が **Proposer** となる場合は，初期に外部から与えられる初期資本1のうち，戦略値 $S_i=(p_i, q_i)$ に応じて手元に $1-p_i$ を残し，対戦相手の **Responder** エージェント j に p_i をオファーする．もしオファー p_i が **Responder** の戦略値 $S_j=(p_j, q_j)$ で定義される許容額 q_j よりも大きければゲームは成立し，**Proposer** は $1-p_i$ ，**Responder** は **Proposer** のオファー p_i を得る．もしオファー p_i が許容額 q_j よりも小さければそのゲームは不成立となって初期資本は没収され，両者利得は0となる．このゲームを **Proposer** と **Responder** の役回りを交代し2回行う．エージェント i が j と対戦して得る利得をまとめると以下になる．

$$\pi_{ij} = \begin{cases} 1 - p_i + p_j, & (p_i \geq q_j \text{ and } p_j \geq q_i) \\ 1 - p_i, & (p_i \geq q_j \text{ and } p_j < q_i) \\ p_j, & (p_i < q_j \text{ and } p_j \geq q_i) \\ 0, & (p_i < q_j \text{ and } p_j < q_i) \end{cases} \quad (2.8.1)$$

このように Ultimatum Game は 2×2 ゲームとは異なり、利得表で利得を表すことは困難であるため、後退帰納法を用いてナッシュ均衡（部分ゲーム完全均衡）を決定することとなる。この方法では、ゲームの最後の行動を考え、行動をとったプレイヤーの効用を最大化する行動は何であるべきかを決定する。次に最後のプレイヤーが（彼女にとって最適な）行動をとったと仮定した場合、最後から二番目の行動を考え、最後から二番目のプレイヤーの効用を最大化する行動を選択する。この過程をゲームの最初の手番まで継続する。こうして残った戦略はすべて部分ゲーム完全均衡点である。

Ultimatum Game の後退帰納法を用いて導出される解は、 $S = (p, q) = (0, 0)$ すなわち他者へのオファーを全く行わず、相手からのオファーはたとえどのように小さな額であっても受け入れるという戦略である。Ultimatum Game では、相手のオファーを拒否した場合双方の利得が 0 となる。合理的エージェントであればそういった戦略を採るとは考えられず、有限であればどのようなオファーでも受け入れる $q = 0$ という戦略（またはそれに限りなく近い戦略）を採る。相手も同様の考え方で同様の戦略を採っているであろうと推量し、結果的に両エージェントは手元に最も多くの資本を残置することができる $p = 0$ の戦略を採るという理由による。これはナッシュ均衡であると同時にパレート最適であるため、Ultimatum Game には何らジレンマ構造は内包されていないといえる。

2.8.2 アダプティブダイナミクス

連続値属性の漸進的進化ダイナミクスについては、アダプティブダイナミクス (adaptive dynamics) を適用した研究[91][92]が行われており、このダイナミクスを用いることで従来の一方向的な進化に加え、分裂的な進化ダイナミクス (evolutionary branching) が可能であることが分かってきている。ここでは、アダプティブダイナミクスを用いた進化の記述を説明する。

アダプティブダイナミクスでは、毎時間ステップ各エージェント（レジデント）は同一の戦略を用いる均一属性の母集団を形成していると仮定する。もちろんエージェントが用いる戦略は連続値属性である。この時、レジデントに対して希少な変異個体（ミュータント）の侵入可能性を定義することで進化のダイナミクスを記述することが出来る。

まず、レジデントが用いる戦略 S を p_1, \dots, p_n の様に n 個定義し、 $S = (p_1, \dots, p_n)$ と書き表すことにする。 $n \rightarrow \infty$ のとき、レジデントは p_1 から p_n の範囲で定義される連続値属性の戦略

を用いていると近似できる．同様に，ミュータントが用いる戦略を $S' = (p'_1, \dots, p'_n)$ と定義する．

ミュータントがレジデント集団に侵入できる可能性を侵入適応度 $E(S', S)$ と呼び，レジデントの適応度 φ_x とミュータント適応度 φ_y の差によって

$$E(S', S) = \varphi_y - \varphi_x \quad (2.8.2)$$

の様に表される．各々の適応度の定義は想定する状況によってさまざまである．

(2.8.2) 式を時間微分して求める侵入適応度の勾配ベクトル場は

$$\dot{p}_i = \left. \frac{\partial E(S', S)}{\partial p'_i} \right|_{S' \rightarrow S}, \quad i=1, \dots, n \quad (2.8.3)$$

となり，これをアダプティブダイナミクスと呼ぶ．この勾配ベクトルを辿ることによって均衡点を探ることが出来る．また，勾配ベクトルが消失する点が存在し，その点は特異戦略点と呼ばれる．この点に収束した場合は分裂的な進化ダイナミクスが生起している場合もあり，さらに高次の項を求めることで分岐したシナリオを追うことが可能であるが，本論ではそれに対する説明は省略する．

2.8.3. Ultimatum Game でのアダプティブダイナミクス

Ultimatum Game をアダプティブダイナミクスによって記述していく．Ultimatum Game における侵入適応度は対戦によってエージェントが獲得した利得とする．(2.8.1)式の利得表を(2.8.3)式の表記と対応するように書き直すと

$$E(S', S) = \begin{cases} 1 - p' + p, & (p \geq q' \text{ and } p' \geq q) \\ 1 - p', & (p < q' \text{ and } p' \geq q) \\ p, & (p \geq q' \text{ and } p' < q) \\ 0, & (p < q' \text{ and } p' < q) \end{cases} \quad (2.8.4)$$

式(2.8.3)と利得表(2.8.4)から，Ultimatum Game のアダプティブダイナミクスは

$$\begin{aligned} \dot{p} &= \begin{cases} -1 & \text{if } p > q \\ 0 & \text{if } p < q \end{cases} \\ \dot{q} &= 0 \end{aligned} \quad (2.8.5)$$

となる．すなわち， $p > q$ のとき， p は $p = q$ となるまで減少し続け， $p < q$ のとき，系は変化せず，系へ影響を及ぼさなくなつて，微分が不可能となり，進化の記述はここで終了してしまう．

ここで従来のアダプティブダイナミクスを一般化し、ランダムドリフトの影響をパラメータ ε を導入して考慮することでこの問題を解決することが出来る。まずエージェントの2次元戦略値 $p \times q$ はそれぞれ $\left[p - \frac{\varepsilon}{2}, p + \frac{\varepsilon}{2}\right] \times \left[q - \frac{\varepsilon}{2}, q + \frac{\varepsilon}{2}\right]$ の一様分布で与える。その時の戦略 $S_2=(p_2, q_2)$ のエージェントが獲得する期待利得は,

$$E(S_2, \bar{S}) = \frac{1}{\varepsilon^2} \int_{p-\frac{\varepsilon}{2}}^{p+\frac{\varepsilon}{2}} \int_{q-\frac{\varepsilon}{2}}^{q+\frac{\varepsilon}{2}} (1-p_2)H(p_2 - q_1) + p_1H(p_1 - q_2) dq_1 dp_1 \quad (2.8.6)$$

となる。ここで $H(x)$ はヘヴィサイド関数であり、 $x>0$ のとき $H(x) = 1$ 、 $x<0$ のとき $H(x) = -1$ である。式(2.8.6)を偏微分し、勾配ベクトル場を求めると、

$$\left. \frac{\partial}{\partial p_2} E(S_2, \bar{S}) \right|_{S_2 \rightarrow S} = \frac{1}{\varepsilon} \int_{q-\frac{\varepsilon}{2}}^{q+\frac{\varepsilon}{2}} -H(p - q_1) + (1-p)\delta(p - q_1) dq_1 \quad (2.8.7)$$

$$\left. \frac{\partial}{\partial q_2} E(S_2, \bar{S}) \right|_{S_2 \rightarrow S} = \frac{1}{\varepsilon} \int_{p-\frac{\varepsilon}{2}}^{p+\frac{\varepsilon}{2}} -q\delta(p_1 - q) dp_1 \quad (2.8.8)$$

となる。ここで、 $\delta(x)$ はディラックデルタ関数であり、 $x \neq 0$ のとき $\delta(x) = 0$ 、 $x=0$ のとき、 $\delta(x) = \infty$ である。上式について詳細に場合分けを行うと、

(i) $p > q + \frac{\varepsilon}{2}$ の場合、

$$\begin{aligned} \left. \frac{\partial}{\partial p_2} E(S_2, \bar{S}) \right|_{S_2 \rightarrow S} &= -1 \\ \left. \frac{\partial}{\partial p_2} E(S_2, \bar{S}) \right|_{S_2 \rightarrow S} &= 0 \end{aligned} \quad (2.8.9)$$

(ii) $p \in \left[q - \frac{\varepsilon}{2}, q + \frac{\varepsilon}{2}\right]$ の場合、

$$\begin{aligned} \left. \frac{\partial}{\partial p_2} E(S_2, \bar{S}) \right|_{S_2 \rightarrow S} \left. \frac{\partial}{\partial p_2} E(S_2, \bar{S}) \right|_{S_2 \rightarrow S} &= \frac{1-p}{\varepsilon} - \frac{1}{\varepsilon} \left(p - q + \frac{\varepsilon}{2}\right) \\ \left. \frac{\partial}{\partial q_2} E(S_2, \bar{S}) \right|_{S_2 \rightarrow S} &\left(p - q + \frac{\varepsilon}{2}\right) \\ \left. \frac{\partial}{\partial q_2} E(S_2, \bar{S}) \right|_{S_2 \rightarrow S} &= -\frac{q}{\varepsilon} \end{aligned} \quad (2.8.10)$$

(iii) $p < q - \frac{\varepsilon}{2}$ の場合、

$$\begin{aligned}\frac{\partial}{\partial p_2} E(S_2, \bar{S}) \Big|_{S_2 \rightarrow S} &= 0 \\ \frac{\partial}{\partial q_2} E(S_2, \bar{S}) \Big|_{S_2 \rightarrow S} &= 0\end{aligned}$$

(2.8.11)

となり，式(2.8.5)を用いて式変形すると，

$$\dot{p} = \begin{cases} -1 & \text{if } p > q + \frac{\varepsilon}{2} \\ 1 - p + O(\varepsilon) & \text{if } p \in \left[q - \frac{\varepsilon}{2}, q + \frac{\varepsilon}{2} \right] \\ 0 & \text{if } p < q - \frac{\varepsilon}{2} \end{cases}$$

$$\dot{q} = \begin{cases} 0 & \text{if } p > q + \frac{\varepsilon}{2} \\ -q & \text{if } p \in \left[q - \frac{\varepsilon}{2}, q + \frac{\varepsilon}{2} \right] \\ 0 & \text{if } p < q - \frac{\varepsilon}{2} \end{cases}$$

(2.8.12)

となる．上式から， $q < p - \frac{\varepsilon}{2}$ のとき， p は減少し続け， q は一定値を保ち，徐々に平均オファークは $p \in \left[q - \frac{\varepsilon}{2}, q + \frac{\varepsilon}{2} \right]$ の範囲に吸引されていく．しかし p がこの閾値内に入れば直ちに適応ダイナミクスによって p は増加し， q は減少するため， $p = q + \frac{\varepsilon}{2}$ 上で振動を繰り返していくことになる．前述のとおり，この状況下では q が増加することはないため，最終的には $S=(0,0)$ へと収束することになる．

3. Ultimatum Game におけるネットワークと戦略適応の共進化モデルに基づく社会的協調の自己組織化

3.1. 研究背景と着眼点

近年は、ネットワーク互惠に対して、どのような付加機構を加えると協調が更に強固に早発されるについて関心が集まっている[3]. 本稿では、相互作用する相手が常に固定されている静的ネットワークの代わりに動的ネットワークを導入し、その影響を考察する. 人や生物の繋がりは生涯を通じて常に一定とは限らないため、ネットワークの可変性を考慮することは妥当な仮定である. ネットワークの進化を規定するにあたって、Zimmermann ら[6]や Tanimoto ら[7]に代表されるように、自分と隣人の戦略の組み合わせや、利得の大小比較によってリンクの切断を決定するのが一般的である. 本論では、これらの既往研究とは視点を変え、戦略の同類性に注目する. 具体的には、戦略が同類である者同士のリンクを維持し、そうでないものとのリンクを切断するものである. この提案ネットワーク進化ルールによって同類的な繋がりを促進することが、囚人のジレンマゲームにおける協調の進化にどのような影響を及ぼすのかを検討する.

3.2. 研究概要

本研究では、Ultimatum Game (UG) を用い、Gao らの報告[93]を基礎にこれを拡張した戦略と topology の共進化モデルを用いる. 社会を構成する N エージェントは、グラフ topology で規定される隣人集合を有する. 各時間ステップで、各エージェントは全隣人とゲーム対戦をした後、ランダムに focal とその隣人 1 人 (すなわち 1 本のリンク) が選ばれ w の確率でネットワーク適応し、さもなくば $1-w$ の確率で戦略適応を行う. すなわち、エージェントの適応は asynchronous に行われる.

さらに本研究では、パラメーター b を導入することによって、従来の UG モデルを拡張する. 拡張モデルでは、自分がオファー p_i を行い、相手に承諾された場合、相手は bp_i の利得を獲得する. この拡張モデルを用いることで、均衡点に応じて異なる利得構造を持たせることが出来、社会効率という指標で均衡戦略値の評価を可能にすることに加え、切断プロトコルが系に与える影響力を計量することが可能となる. 本拡張パラメーターによって定義された系では、自身がより多くの資本を拠出し、一時的には個人の保有資産が減少する状況にな

ればなるほど社会全体としては大きな利益が生み出される状況となり，社会ジレンマ構造がある状況であるといえる．

一般的には，UG は連続値属性の戦略パラメーターを定義して実験が行われるが，本研究では定義される戦略値を離散的にすることで結果に差異生じるかも検証する．

3.3. モデル

3.3.1. 拡張 Ultimatum Game

既往の研究で一般的に用いられている UG モデルは，均衡時における社会全体の利得 Π_{total} は，平均場近似が適用出来て，ネットワーク平均次数を $\langle k \rangle$ ，均衡時の社会平均オファー \bar{p}^* により，

$$\begin{aligned}\Pi_{total} &= \langle k \rangle N (1 - \bar{p}^* + \bar{p}^*) \\ &= \langle k \rangle N \\ &= Const.\end{aligned}\tag{3.3.1}$$

で与えられる．ここで， N は系の総粒子数を表す．この場合，Proposer がオファーする額と手元に残置する額との合算はどのようなオファー額の均衡に至ったとしても 1 になるので，ネットワーク適応により進化した均衡時の次数分布と社会利得分布とは同形となる．現行のモデルでは，式(3.3.1)に表わされている通り，どのようなオファー額の均衡に至ろうとも社会全体の利得は初期に決定した平均次数と粒子数による定数になる．そこで本研究では社会利得という指標での評価を可能にするため，パラメーター b を導入してモデルを拡張し，利得表は以下のように再定義した．

$$\pi_{ij} = \begin{cases} 1 - p_i + bp_j, (p_i \geq q_j \text{かつ} p_j \geq q_i) \\ 1 - p_i, (p_i \geq q_j \text{かつ} p_j < q_i) \\ bp_j, (p_i < q_j \text{かつ} p_j \geq q_i) \\ 0, (p_i < q_j \text{かつ} p_j < q_i) \end{cases}\tag{3.3.2}$$

拡張モデルにおいて先と同様に平均場近似を適用して社会全体の利得 Π_{total} を計算すると， $b \neq 1.0$ の場合，

$$\begin{aligned}\Pi_{total} &= \langle k \rangle N (1 - \bar{p}^* + b\bar{p}^*) \\ &= \langle k \rangle N \{1 - (1 - b)\bar{p}^*\end{aligned}\tag{3.3.3}$$

となり， 喩え全エージェントのオファーが成立している均衡であっても， エージェント毎に異なる p により各自の利得は異なってくるので， 次数分布と利得分布の形状は異なる． この拡張によって $b > 1.0$ のとき， 最終的な均衡点におけるオファーが大きいほど社会全体の利得， すなわち社会効率が大きいと云え， オファー大の均衡点に至るほど効率的であると明確に評価することができるようになる．

本研究では， $b = 1.0$ (デフォルトモデル) と $b = 2.0$ を仮定した．

3.3.2. 離散戦略

離散戦略のとき， エージェントが用いることが出来る戦略値は (p, q) , $(0, 0)$, $(0.5, 0)$, $(1, 0)$, $(0, 0.5)$, $(0.5, 0.5)$, $(1, 0.5)$, $(0, 1)$, $(0.5, 1)$, $(1, 1)$ の 9 戦略のみに限る． その他の枠組みに違いはない．

離散戦略の場合も， $b = 1.0$ (デフォルトモデル) と $b = 2.0$ を仮定した．

3.3.3. ネットワーク適応における切断プロトコル

focal となるエージェント i とその隣人 j をランダム選択し， 以下に述べる 5 つの切断プロトコルでリンクの切断， 保持を決定する． 切断を行ったエージェントは， ネットワーク上の全エージェントからランダム選択したエージェント k に新規のリンクを架ける (但し自己接続， 2重接続を禁ずる)． また， 隣人が存在せずゲームが出来ない Lonely エージェントの出現を極力抑えるため， focal か隣人どちらかが次数 1 の場合は次数が 1 でないエージェントが無条件に切断される (但し両者次数 1 の場合は切断プロトコルに則りどちらかが切断される)． 共進化モデルに基づく UG では， エージェントの利得は保持するリンク数により決まるので， エージェントにはリンク切断後の再リンク権を確保するような淘汰圧が作用することになる．

- Responder 基準[93]

$p_i \geq q_j$ かつ $p_j \geq q_i$ のとき， i - j リンクは保持される．

$p_i \geq q_j$ かつ $p_j < q_i$ のとき， i は j とのリンクを切断し， 新たな相手 k とランダム接続する．

$p_i \geq q_i$ かつ $p_i < q_j$ のとき， j は i とのリンクを切断し， 新たな相手 k とランダム接続する．

$p_i < p_j$ かつ $p_j < q_i$ のとき， i - j リンクは崩壊． i か j のどちらか一方が新たな相手 k とランダム接続する．

- Proposer 基準切断

$p_i \geq q_j$ かつ $p_j \geq q_i$ のとき, i - j リンクは保持される.

$p_i \geq q_j$ かつ $p_j < q_i$ のとき, j は i とのリンクを切断し, 新たな相手 k とランダム接続する.

$p_j \geq q_i$ かつ $p_i < q_j$ のとき, i は j とのリンクを切断し, 新たな相手 k とランダム接続する.

$p_i < p_j$ かつ $p_j < q_i$ のとき, i - j リンクは崩壊. i か j のどちらか一方が新たな相手 k とランダム接続する.

- オファー比率基準

$p_i \geq q_j$ かつ $p_j \geq q_i$ のとき, i - j リンクは保持される.

$p_i \geq q_j$ かつ $p_j \geq q_i$ 以外 のとき, 確率 $p_i/(p_i+p_j)$ で i は j とのリンクを切断し, さもなくば j が i とのリンクを切断する. 切断を行ったエージェントは新たな相手 k とランダム接続する.

- 寛大さ比率基準

$p_i \geq q_j$ かつ $p_j \geq q_i$ のとき, i - j リンクは保持される.

$p_i \geq q_j$ かつ $p_j \geq q_i$ 以外 のとき, 確率 $(p_i - q_i)/(p_i - q_i + p_j - q_j)$ で i は j とのリンクを切断し, さもなくば j が i とのリンクを切断する. 切断を行ったエージェントは新たな相手 k とランダム接続する.

- インセンティブ比率基準

$p_i \geq q_j$ かつ $p_j \geq q_i$ のとき, i - j リンクは保持される.

$p_i \geq q_j$ かつ $p_j \geq q_i$ 以外 のとき, 確率 $(2p_i - q_i - p_j)/(p_i + p_j - q_i - q_j)$ で i は j とのリンクを切断し, さもなくば j が i とのリンクを切断する. 切断を行ったエージェントは新たな相手 k とランダム接続する.

3.3.4. 各切断プロトコルの詳細

3.3.4.1. 確定的切断プロトコル

本研究では, 切断プロトコルの性質から各切断プロトコルを確定的な切断プロトコルと確率的な切断プロトコルの2種類何れかに分類する. Responder 基準切断プロトコル

(Responder Rule) と, Proposer 基準切断プロトコル (Proposer Rule) は, オファーの大小に関わらず, オファーの拒諾に応じて確定的に切断を行うことから, 確定的切断プロトコルと分類する. また, 切断が行われるのは両者のオファーが一方でも不成立となった場合であり, 両者のオファーが共に成立している場合はリンクが保持されるため, 3.3.3.と 3.3.4.での切断権に関する説明は, すべてリンク保持となる状況以外に対してであることを先に述べて

おく。Responder 基準では、オファーを受け取らなかったエージェントにリンク切断権を与えている。これは、相手オファーが自己受容閾値より小さいと不満を感じるとの前提に依拠しており、なぜそのようにオファーが小さいのかという考え方に基づいている。Proposer 基準は、Responder 基準とは逆で、相手のオファー受容閾値が自分のオファーより大きいと不満を感じるとの前提に依るものであり、なぜオファーを受け取らないのかという考え方である。こういった切断方法も現実でのさまざまな交渉場面において十分に考え得ることのできる状況である。また、両者のオファーが共に不成立となった場合は、共に不満を抱いたと見做して $i-j$ のリンクは崩壊し、どちらか一方がランダムに接続する条件となっている。

3.3.4.2. 確率的切断プロトコル

確率的切断プロトコルは、リンク保持となった状況以外において、各エージェントが保有する戦略値の大小を加味して切断、接続を行うプロトコルである。確定的な切断プロトコルよりもさらに厳密な切断を行うことができる。オファー比率基準 (Offer Rule) では、各エージェントのオファーの大きさの比率に応じた確率 $p_i / (p_i + p_j)$ で i 、さもなくば j に切断権を付与するものである。このプロトコルは、オファーが大きいエージェントほど切断権を与えられるべきであるとの考えに依拠している。寛大さ比率基準 (Generous Rule) では、確率 $(p_i - q_i) / (p_i - q_i + p_j - q_j)$ で i 、さもなくば j に切断権を付与する。これは、オファー p と許容閾値 q の差が大きい、すなわち他者への大きなオファーを行いどのようなオファーでも受け入れる寛大なエージェントほど切断権を手に入れるべきであるという考えに依拠している。インセンティブ比率基準 (Incentive Rule) は、確率 $(2p_i - q_i - p_j) / (p_i + p_j - q_i - q_j)$ で i 、さもなくば j に切断権を付与している。インセンティブ比率基準の考え方は、本来対戦する2人のエージェントは明確に相手の許容閾値を知ることは出来ず、分かるのはオファーの拒諾のみである筈であるという前提に基づき、相手の受容閾値を自身のオファー額と置き換えて自分の寛大さを $p_i - q_i$ 、相手の寛大さを $p_j - p_i$ と評価し、その差 $2p_i - q_i - p_j$ を i のインセンティブ、同様に $2p_j - q_j - p_i$ を j のインセンティブと定義してインセンティブの比率に応じた確率で切断権を付与する。本切断プロトコルは、本来エージェントの知り得ることのできないグローバルな情報をゲーム構造に組み込むことは不適切である可能性があるため、よりローカルな情報のみでゲームを展開させることを目的としている。

3.3.5. シミュレーション設定

エージェント集団サイズは $N = 1600$ ，初期ネットワークは平均次数 $\langle k \rangle = 12$ の E-R ランダムグラフ，戦略適応は Pairwise-Liner[93]を用いる．戦略適応時コピーしてきた戦略にはコピーノイズ $[-\delta, +\delta]$ が p_j と q_j にそれぞれ別々に印加される． $\delta = 0.005$ とし，100 試行のアンサンブル平均を解析対象とする．但し，離散戦略での実験の場合は戦略適応時にコピーノイズを印加しない．その代わりに，各世代突然変異確率 μ でランダムに選ばれたエージェントが 9 戦略の何れかに変異する．どちらの場合も $b = 1$ と $b = 2$ の 2 パターンを仮定する．本研究では $\mu = 0.001$ とした．各ステップにおいてフォーカルエージェントと隣人エージェントがランダムに一人ずつ選ばれ，確率 w でネットワーク適応を行い，確率 $1 - w$ で戦略適応を行う．

疑似均衡に要するステップ数の判定は集団サイズとノイズに対して敏感である．図 3.3.1 に $b = 1.0$ ， $w = 0.1$ ，Responder 基準におけるあるエピソードを示す．本論では戦略値が十分に均衡に達したとみなされた時点で 1 試行シミュレーションを打ち切る．

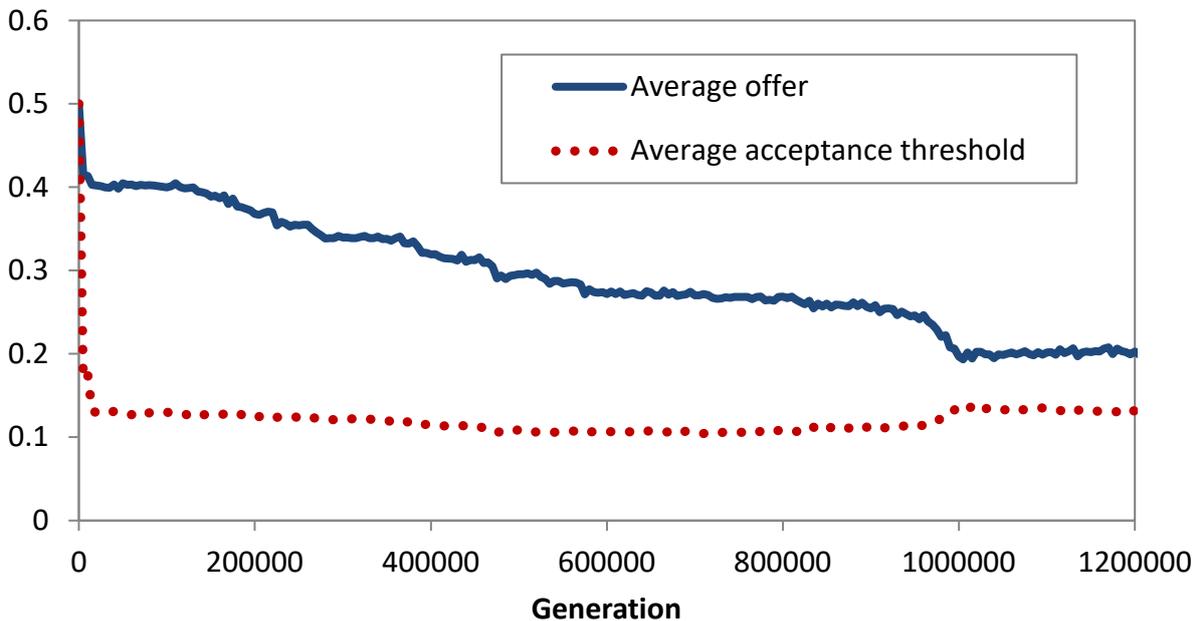


図 3.3.1 Responder 基準切断プロトコル， $w = 0.1$ のときのある 1 試行における社会平均オファー額と社会平均許容額の時系列変化．

3.4. 結果と考察

3.4.1. 次数分布

図 3.4.1 に $b = 1.0$ の場合におけるそれぞれの切断プロトコルごとの次数分布を示す. 1つの切断プロトコルの図中にネットワーク適応の速度パラメーター w は $w = 0, 0.3, 0.7, 0.9$ の 3 パターンを掲載している.

ネットワークは, w が大きくなるに従って初期の Poisson 分布からよりヘテロな次数分布へと進化する (Proposer 基準を除き, その分布特性は切断プロトコルや b には大きく依存しなかった). 2×2 ゲームにおける共進化モデル同様, リンク切断という Game Exit オプション[21]の行使により, よりヘテロなネットワークへと社会を適応させることで, 固定ネットワークゲーム ($w = 0$) より高い公平性と協調性の進化が可能となる. Proposer 基準を除けば, ネットワーク適応確率 w が十分に大きければ, 固定グラフ上のゲームよりも大きな p へと進化しており, より強い互惠関係が創発していることがわかる. これは, 2×2 ゲームにおける検討で Tanimoto や Santos[94][95]が指摘しているように, 戦略適応に対してネットワーク適応の時間スケールが短くなることで, 協調進化に利していることによる. つまり, focal エージェントはリンク切断プロトコルで規定された不都合な対戦相手との関係を即座に解消出来るため, p の大きな戦略の生き残りが可能になる.

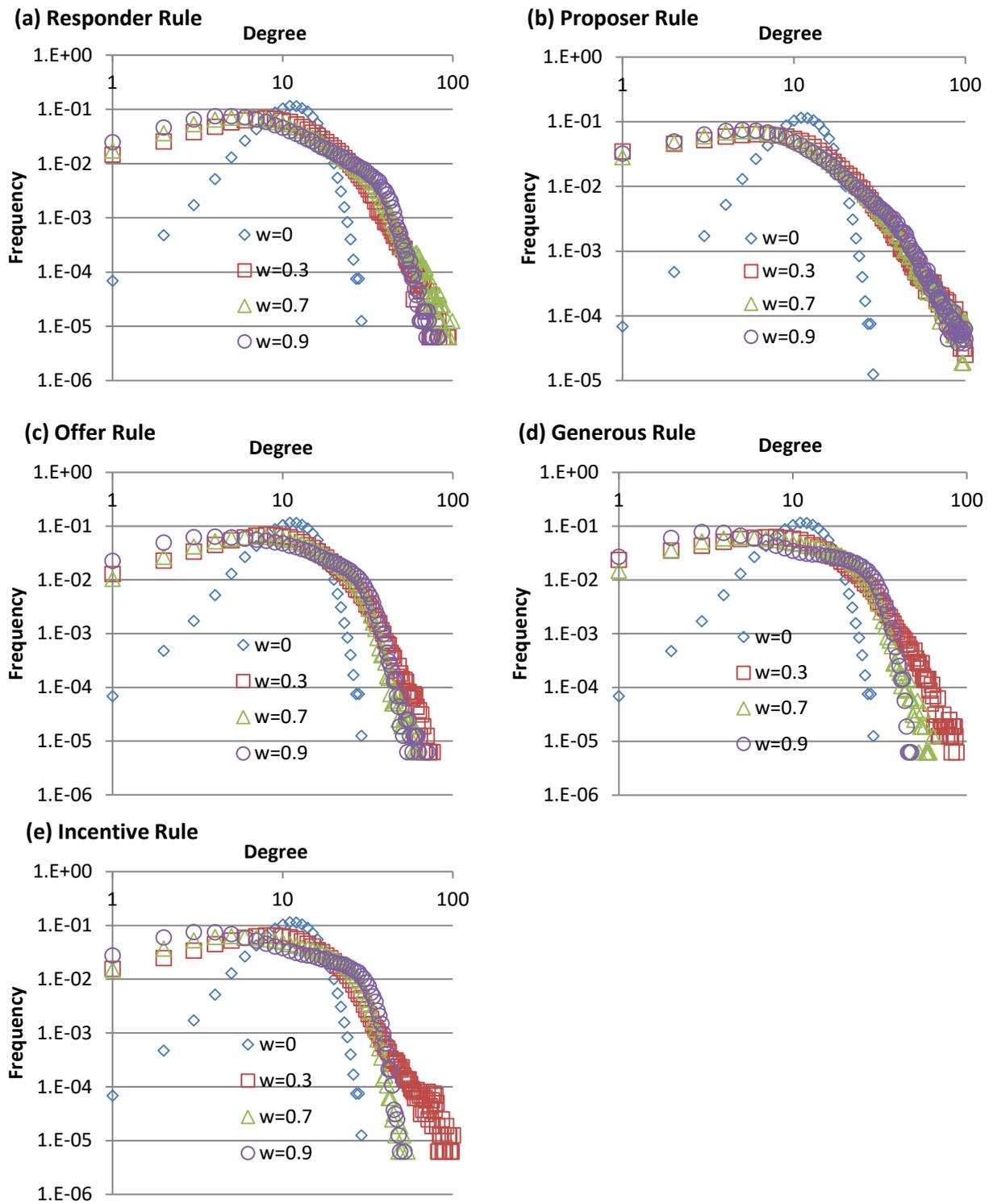


図 3.4.1 $b = 1.0$ の場合における各切断プロトコルごとの次数分布. 1つの切断プロトコルで $w = 0, 0.3, 0.7, 0.9$ の 4 パターンを掲載している.

3.4.2. 戦略値進化の考察

3.4.2.1. $b = 1.0$ の場合

図 3.4.2 に $b = 1.0$ の場合での各切断プロトコルにおける均衡時の社会平均戦略 p , q とネットワーク適応の速度パラメーター w との関係を示す. 図 3.4.3 に各プロトコルにおける p と q の均衡値を w の小さい順にラインで結び, 背景に p と q の均衡値で決まる社会利得をコンターで示している. 既述のように $b = 1.0$ では社会効率性は p の大小によらないが, $b = 2.0$ では大きな p ほど社会平均利得は大きくなる.

$b = 1.0$ の場合, w 大のとき Responder 基準では, p は公平な配分 0.5 を大幅に上回った 0.9 周辺のかなり大きな値に, それに伴ってオファー受容閾値 q は 0.4 付近で均衡している. これは初期戦略分布上限を 0.5 で与えている Gao ら[93]の結果とは異なる.

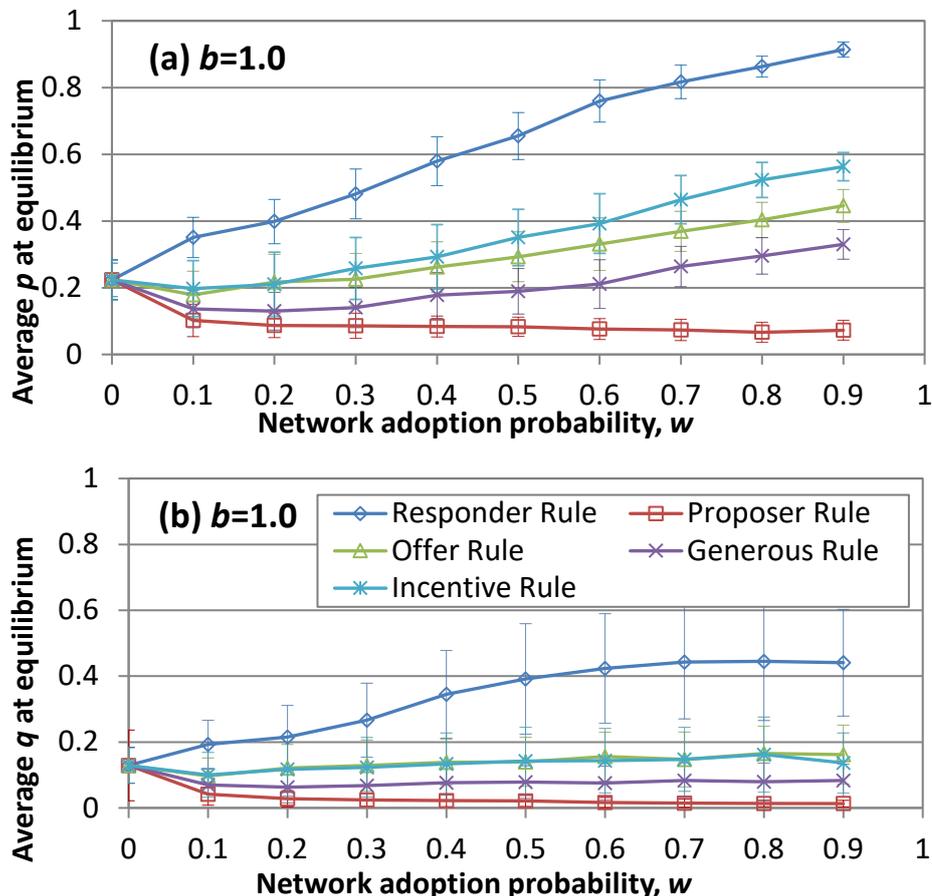


図 3.4.2 それぞれの w における均衡時の(a)社会平均オファーと(b)社会平均許容額.

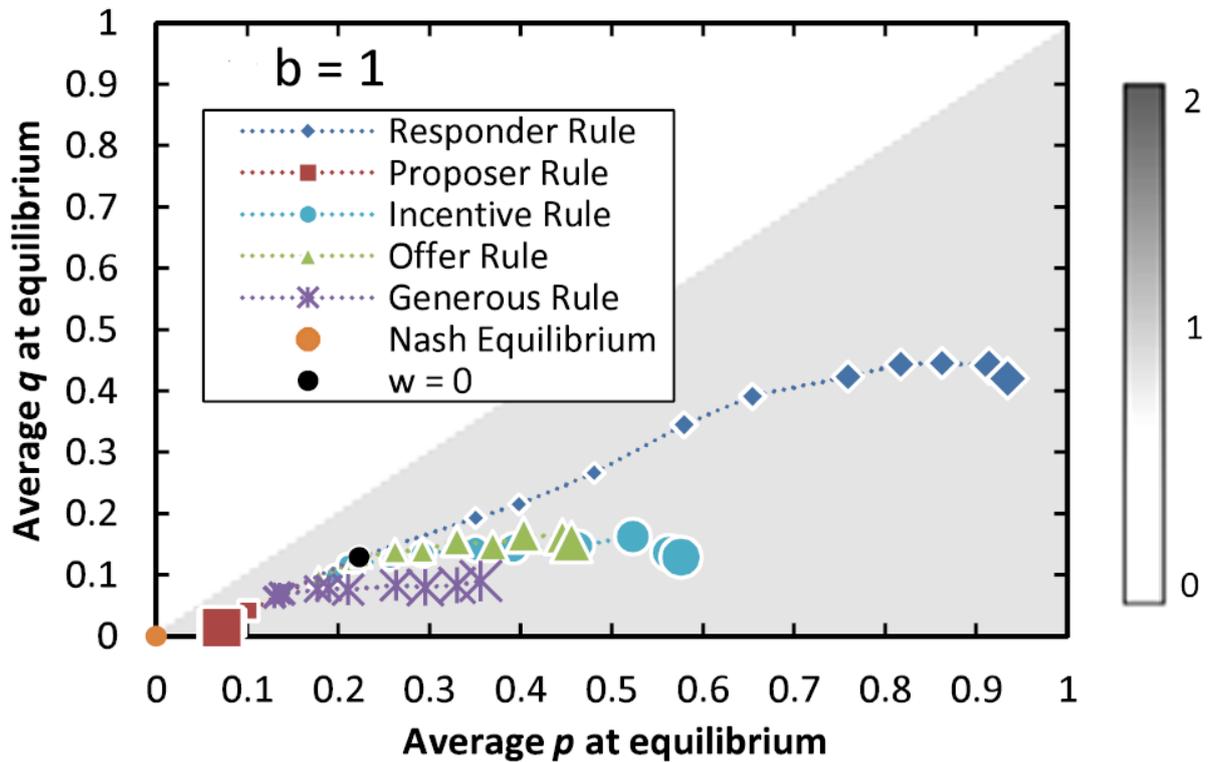


図 3.4.3 均衡時における平均オファー p と平均許容額 q の関係図。プロットが大きくなるほどネットワーク適応確率が大きくなっていることを表す。背景のコンター図は平均場近似を用いて計算された均衡時の社会利得を表す。

Responder 基準では、いかに大きなオファーをしても、相手の許容閾値が大きいためゲーム不成立となれば、オファーの主は確定的に切断される。各エージェントの利得は隣人との全対戦で獲得した利得積算なので、自らの次数を大きくする側、つまり相手に切断されない方向へと社会は進化する。切断されないためには、「気前のよい」ことが適応的になるから、 w が大きくなるに従って過度に大きな p へと進化する。

Proposer 基準では、 p 、 q 共に 0 よりは若干大きい絶対的には小さな値に落ちつく。このプロトコルでは、いかに小さなオファーに対しての拒否であろうとも、ゲームを不成立にしたエージェントのリンクが確定的に切断される。そのため、リンクの切断を免れ、自らの次数を減じさせないためには、相手のいかなるオファーに対して受容する「寛大な」ことが適応的になるため、 q は過度に小さい均衡に至る。相手の許容閾値を満足させることが出来れば敢えて大きなオファーを出す必要がなくなるので、 p が他プロトコルに比し相対的に小さい戦略へと均衡する。

オファー比率基準では、 p が大きいほど高い確率で切断権を手に入れることが出来、自らの次数減少を免れ得るので、 w 大に従って p が大きくなっているが、Responder 基準ほどではない。これは、どんなに大きなオファーをしても相手を満足させることが出来なければ確実に切断されてしまう Responder 基準とは違い、たとえフォーカルのオファーが拒否されたとしてもフォーカルとその対戦相手とのオファーの大きさに応じた確率で切断権を付与するため、Responder 基準ほど極端な均衡とはならないことによる。このように Responder 基準や Proposer 基準のような確定的切断権付与プロトコルでは、被験者実験結果に照らしても [96][97][98][99][100][101]、現実の人間社会ではあり得ないような極端な気前よさ (p 大) や寛大さ (q 小) に均衡してしまうが、確率的プロトコルでは妥当な均衡へと吸引される。

寛大さ比率基準は、オファー比率基準よりも約 0.1 ほど p は小さい。寛大さ比率基準では、切断権の所在が p - q の値に依るため、 q が小さいエージェントも適応的になり得る。Proposer 基準でも述べたが、 q が小さいエージェントに対して敢えて大きな p を提示する必要はないため、オファー比率基準より若干 p , q 共に小さい均衡へと至ったと考えられる。

インセンティブ比率基準では、 i のインセンティブ $2p_i - q_i - p_j$ と j のインセンティブ $2p_j - q_j - p_i$ の比率によって確率的に切断権を付与する。これは自身のオファーを大きくすることで自分のインセンティブを上昇させることができると同時に相手のインセンティブを下げることもできるため、先の 2 つのプロトコルよりも p 大への進化が（但し、Responder 基準よりは低い）促進されたものと考えられる。

以上より、戦略値の最終的な均衡点は共進化における切断プロトコルに大きな影響を受けている事が分かった。

3.4.2.1.1. 不合理切断の生起回数

Responder 基準では、比較的大きなオファーを出しても相手の高いオファー受容閾値により切断されるとの非現実的切断が生じると述べた。ここでは、非現実的切断がどの程度の割合で生起しているのかを見ていく。図 3.4.4 に条件ごとの切断頻度を示す。すべての値は Responder 基準切断プロトコルと比べて何倍かを示している。

$w = 0.9$ のとき、 $p_i \geq 0.5$ のときに切断されるイベント (■) が非現実的切断であると定義すれば、どの切断プロトコルも Responder 基準でのイベント生起回数よりも少なくなっていることがわかる。特に、インセンティブ比率基準によるそのイベント生起回数は

Responder 基準のその 0.35 倍となっており、不合理切断が 1/3 近くまで減少している。

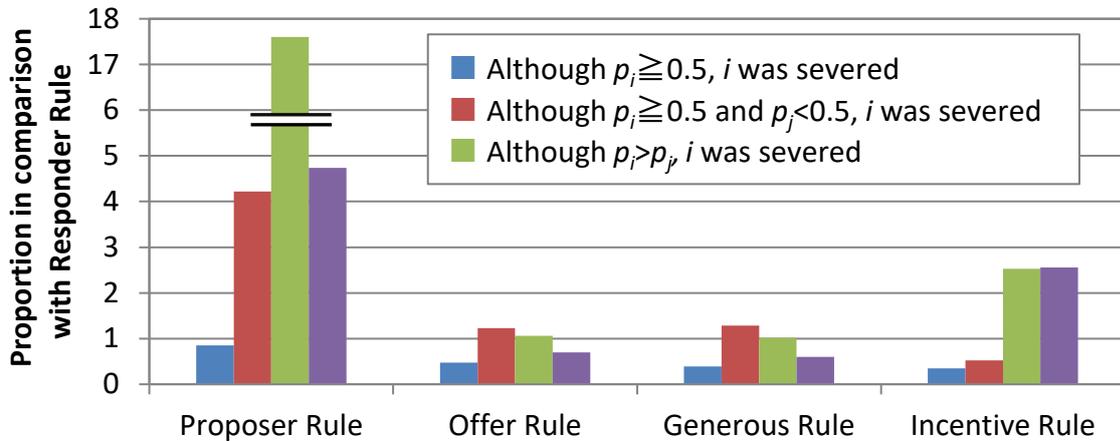


図 3.4.4 $w = 0.9$ における各切断プロトコルでの条件別切断頻度。切断頻度の基準は Responder 基準切断プロトコルとしている。すなわち、1 つの棒グラフの値は Responder 基準切断プロトコルでの値と比べて何倍になっているかを表す。

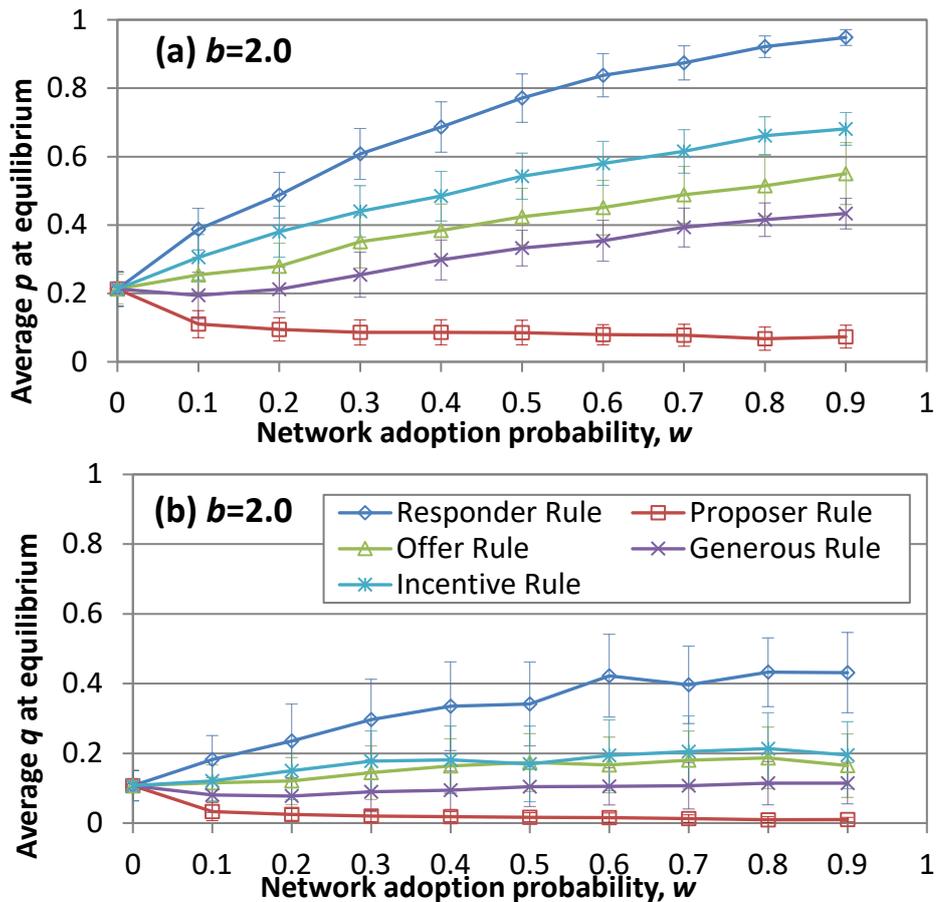


図 3.4.5 $b = 2$ におけるそれぞれの w における均衡時の(a)社会平均オファーと(b)社会平均許容額。

続いて $p_i \geq 0.5$, $p_j < 0.5$ のときに切断されるイベント (■) を見てみよう。Proposer 基準は, Responder 基準に比べて 4 倍以上の生起回数になっている。これは, Proposer 基準ではオファーを受け取らなかったエージェントが確定的に切断されるため, いくら i が公平なオファーを提示しても, i が j の (不公平な) オファーを拒絶したために切断されてしまっていることを示しており, 極めて不自然な切断がかなりの割合で生起していることが窺える。

$p_i > p_j$ のときに切断されるイベント (■) を見ると, Proposer 基準が突出した生起回数を示している。また, Incentive 比率基準も約 2.5 倍の生起回数になっている。しかし, Incentive 比率基準は $p_i \geq 0.5$, $p_j < 0.5$ のときに切断されるイベント (■) の生起回数が少ないため, ほとんどのイベントは $p < 0.5$ のエージェント同士の対戦で生起したものと推測され, 然程不自然ではなく, 寧ろ真つ当な切断イベントが営まれているといえる。

最後に実際に切断されるイベントが生起した割合 (■) を見てみよう。Proposer 基準, Incentive 比率基準は Responder 基準に比べて切断回数が多くなっている。Proposer 基準は p と q の値が非常に近い点で均衡する。シミュレーション時間が長くなるほどコピーノイズの影響が効いて戦略が変動するため, p と q の値が非常に近い Proposer 基準では切断イベントが再生起しやすいためだと考えられる。Incentive 比率基準は, 他の切断プロトコルよりも切断権付与の式が複雑であるため, 1 試行で均衡に至るまでは p と q の分布に多様性が生じやすく, 切断イベントが生起しやすくなっていると推定される。

以上より, 確定的切断プロトコル (Responder 基準, Proposer 基準) に比べて不合理切断イベントが生起する回数が減少していることがわかる。

3.4.2.2. $b = 2.0$ の場合

続いて $b = 2.0$ の結果を切断プロトコルごとに見ていく。図 3.4.5 に $b = 2.0$ の場合での各切断プロトコルにおける均衡時の社会平均戦略 p , q とネットワーク適応の速度パラメーター w との関係を示す。

図 3.4.6 に各プロトコルにおける p と q の均衡値を w の小さい順にラインで結び, 背景に p と q の均衡値で決まる社会利得をコンターで示している。既述のように $b = 1.0$ では社会効率性は p の大小によらないが, $b = 2.0$ では大きな p ほど社会平均利得は大きくなる。

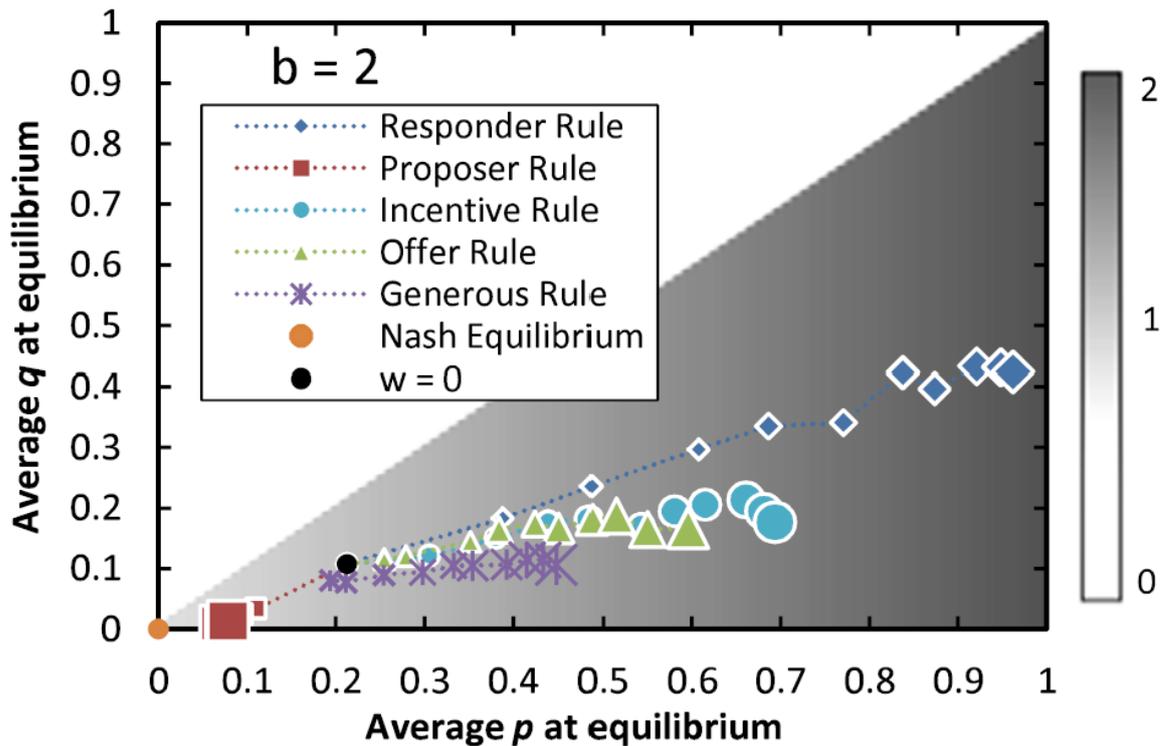


図 3.4.6 均衡時における平均オファー p と平均許容額 q の関係図. プロットが大きくなるほどネットワーク適応確率が大きくなっていることを表す. 背景のコンター図は平均場近似を用いて計算された均衡時の社会利得を表す.

Responder 基準では $b = 2.0$ は, 特に w が小さいとき $b = 1.0$ の場合に比べて若干オファーが大きくなっているが, w が大きくなると大差ない. これは切断プロトコルの特性上, $b = 1.0$ でも既に十分大きな p へと進化しており, 利得構造変容の (b が大きくなることによる) の影響を受ける余地は小さいためだと考えられる.

Proposer 基準の場合, 戦略値は $b = 1.0$ の場合とほぼ同様の結果である. $w = 0$ (●) に比べて $w > 0$ では社会全体の利得が減少する方向へと進化している. ネットワーク適応の速度が速くなると, 相手に切断されてリンクを失うことによる利得減少が大きく影響するようになるため, 公平な Pareto 最適となるオファーを大きくして (相手からも同様のオファーをもらうことを期待して) ジレンマを克服する方向に進化しない. このことは, ネットワーク適応が互惠創発に悪影響していると言うことが出来, 2×2 ゲームにおける共進化モデルと様相が異なる. 共進化によりリンク切断することが, Ultimatum Game 以外のゲーム性を系に付加し, リンク消失による利得減少が, 互惠による利得増大の影響を凌駕し, 結果的に共進化のない固定グラフ上ゲームの均衡より劣位の均衡にしか至り得なくなっている.

オファー比率基準，寛大さ比率基準，インセンティブ比率基準では，どれも $b = 1.0$ に比べて $0.1 \sim 0.2$ 程度 p が大きくなっており，ネットワーク適応速度が大きくなるに従って社会効率が大きくなっている．確定的切断権付与により非現実的切断イベントが起きる Responder 基準を除外すれば，インセンティブ基準が最も公平な Pareto 最適に近いより大きな p に均衡し，強固な互惠関係が進化している．

3.4.3. ネットワークパラメーター

結果として示すネットワークパラメーターは次数相関，クラスター係数，平均パス長，孤立した島の数の4つである．

次数相関，クラスター係数，平均パス長の3パラメーターの特性は第2章で説明したとおりである．孤立した島の数とは，ネットワーク適応によってネットワークが完全に分断されたクラスターがいくつ存在しているかを示す（図 3.4.7）．

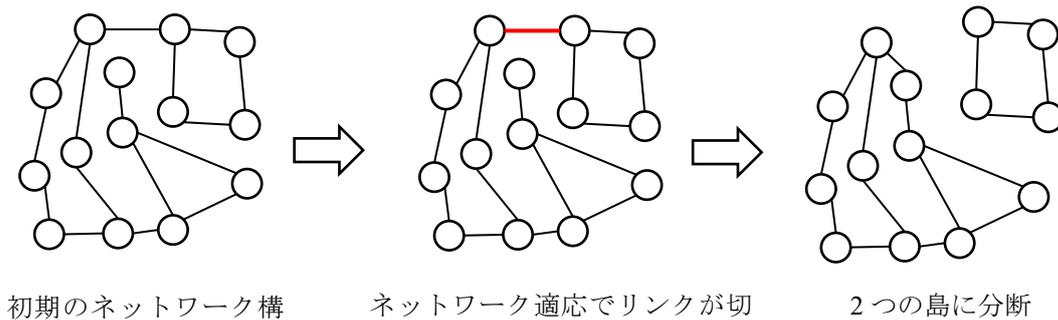


図 3.4.7 ネットワーク適応によって独立した島が出現するプロセス．

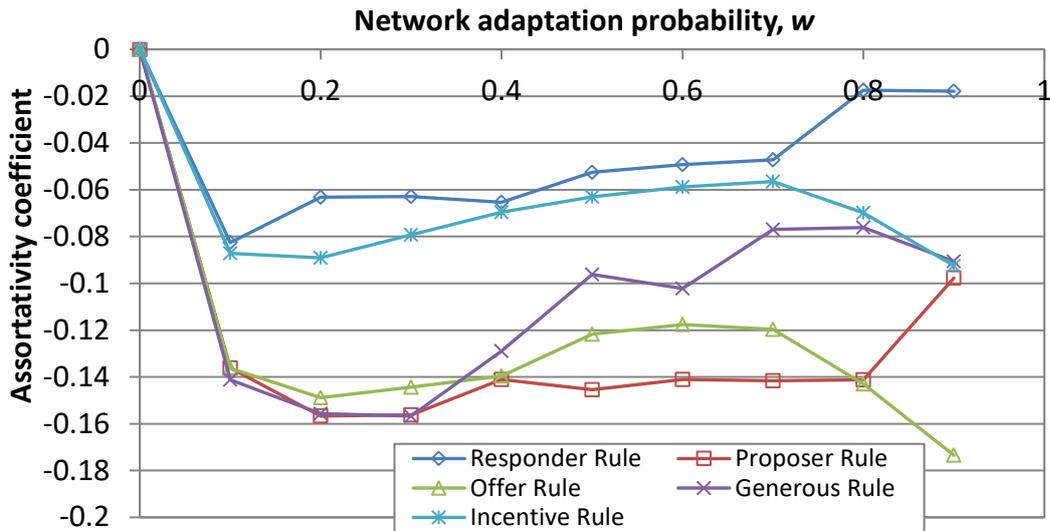


図 3.4.8 各切断プロトコルにおける次数相関とネットワーク適応確率の関係。

図 3.4.8 に各切断プロトコルにおける次数相関とネットワーク適応確率の関係を示す。図を見ると、全領域で次数相関は負となっていることがわかる。つまり次数相関を負の方向へと振ることで協調関係を維持していると云える。ここでいう協調関係とは、ナッシュ均衡ではない有限のオファーを行うことを指し、必ずしも公平なオファー $p=0.5$ を指しているわけではない。次数相関が負だと、次数の小さいエージェントが次数の大きな所謂ハブと呼ばれる存在とネットワークを持っていることになる。本研究では戦略適応方法に **Pairwise Linear** を用いているため、協調的なオファーを行う次数が大きいハブエージェントが協調的な戦略を採れば、ハブに取り付いている次数の小さなエージェントはハブの戦略をコピーし続け、ナッシュ均衡へと吸引されず非常に強固な協調クラスターが維持されるためだと考えられる。仮にハブに取り付いているエージェントの誰かが戦略コピー時のノイズによる影響で（各切断プロトコルによって定義される）利己的な戦略へと進化してしまっても、切断プロトコルによって直ちに切断されるためクラスターがナッシュ均衡側へと吸引されることはない。

図 3.4.9 に各切断プロトコルのクラスター係数を示す。

Proposer 基準を除いて、どの切断プロトコルも初期のクラスター係数から共進化によってやや値が上昇しているが、ほぼ変化はない。**Proposer** 基準は他の切断プロトコルよりも値がやや大きい。

図 3.4.10 に平均パス長を示す。

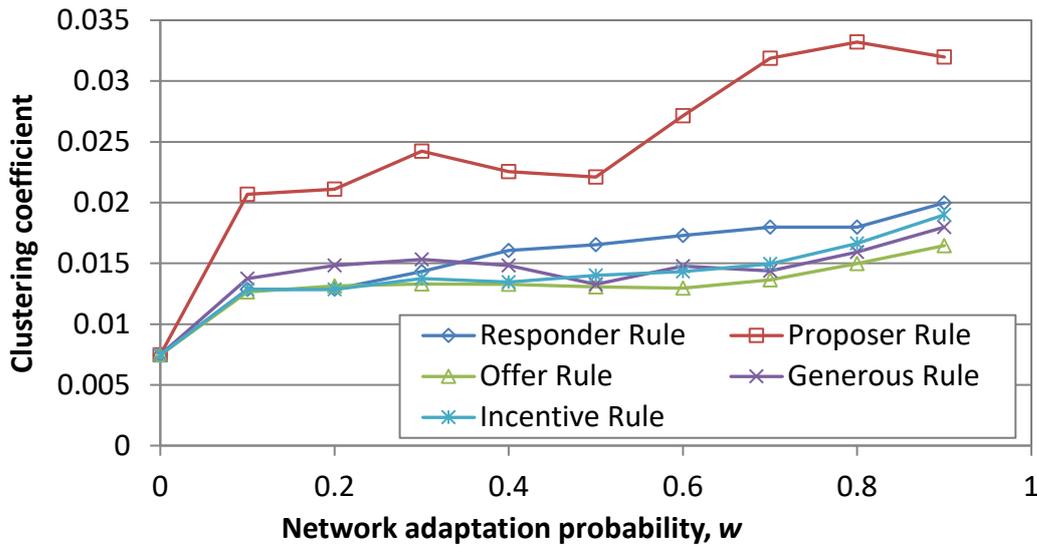


図 3.4.9 各切断プロトコルにおけるクラスター係数とネットワーク適応確率の関係。

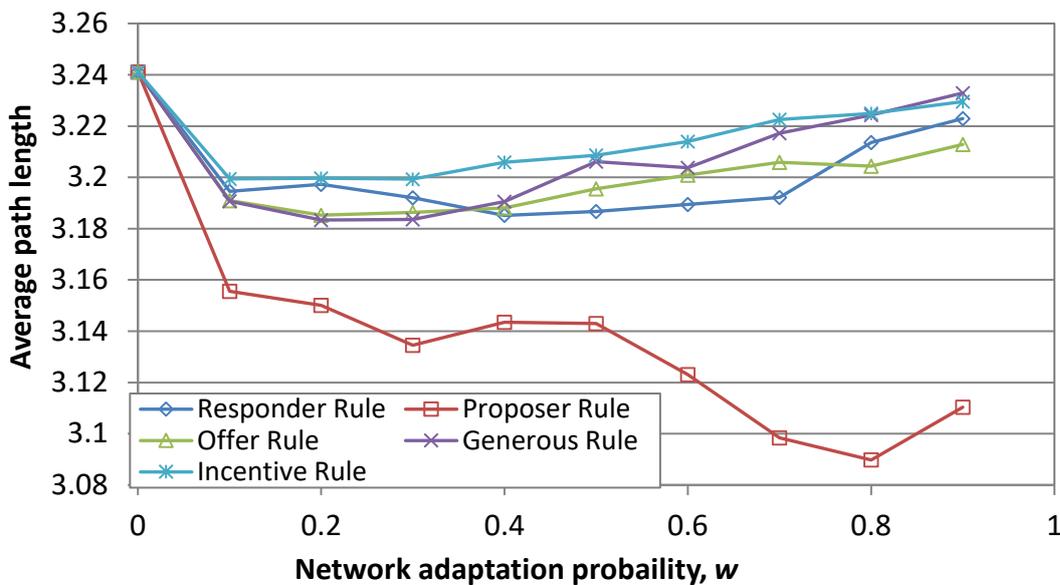


図 3.4.10 各切断プロトコルにおける平均パス長とネットワーク適応確率の関係。

Proposer 基準以外は特徴的な変化は見られない。ここで図 3.4.9 中の Proposer 基準の結果を見てみると、クラスター係数の変化と平均パス長の変化が対応していることがわかる。つまり、Proposer 基準は他の切断プロトコルに比べてフォーカル近傍のエージェントと密接に繋がっており、そのため、平均パス長が短くなっているものと考えられる。

図 3.4.11 に島の数を示す。

これを見ると Responder 基準、Proposer 基準といった確定的切断プロトコルだとネットワーク適応確率が大きくなるにつれて出現する島の数が確率的切断プロトコルよりも明らかに多い。これは確定的切断プロトコルが初期に極端な切断を行うため島が分断されやすくなって

いると考えられる。また、図 3.4.4 に示した切断回数を見ると、Proposer 基準は切断の生起回数が他の切断プロトコルに比べて非常に大きいことが分かり、単独の島が出現しやすくなる要因となっていると考えられる。

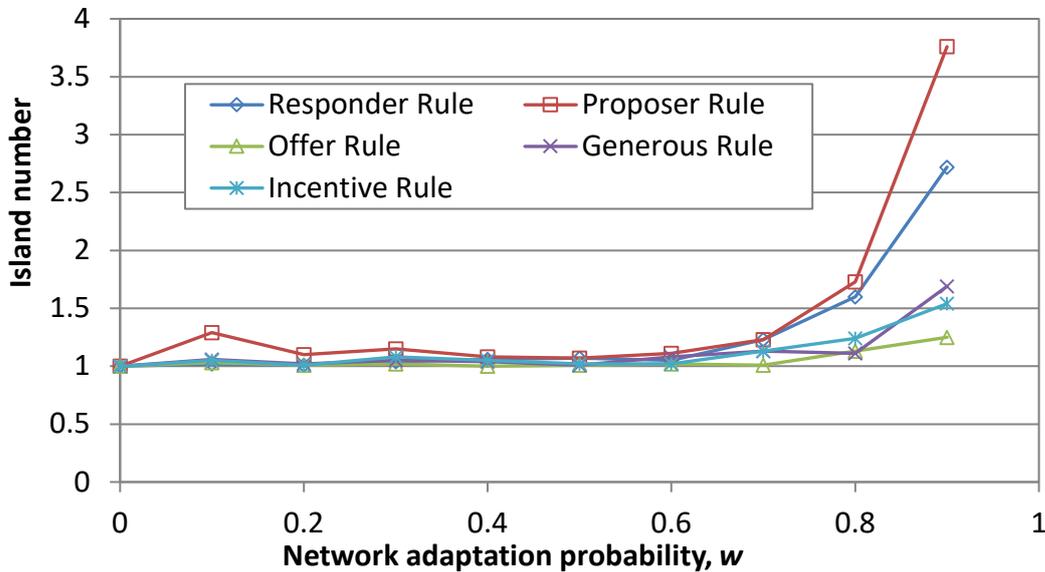


図 3.4.11 各切断プロトコルにおける島の数とネットワーク適応確率の関係。

3.4.4. 離散戦略

離散戦略の場合、各世代における突然変異確率は $\mu = 0.001$ とした。

3.4.4.1. $b = 1.0$ の場合

図 3.4.12 に離散戦略、 $b = 1.0$ のときの社会平均オファーと社会平均許容額を示す。この図を見ると、連続戦略の場合と異なる結果となっていることがわかる。これは、戦略値が離散値で与えられるため、0 や 0.5, 1 といった端点に収束しやすいためであることは容易に想像できる。

Responder 基準、オファー比率基準、インセンティブ比率基準は $w = 0.1 \sim 0.6$ にかけて極めて公平なオファー $p = 0.5$ となっている。それ以上ネットワーク適応速度が大きくなると Responder 基準は $p = 1$ 、インセンティブ比率基準は $p = 0.7$ 付近で均衡し、オファー比率基準は全く変化がなく常に $p = 0.5$ に均衡している。このことから、離散戦略の方がネットワーク適応速度に依存せず公平性が創発するといえる。

Proposer 基準、寛大さ比率基準を見てみると、完全にナッシュ均衡に吸引されている ($w = 0.9$ を除く)。Proposer 基準は先に述べられた通り、その切断プロトコルの特性上 p と q が小さな均衡へと吸引されていくのは尤もらしい結果である。

しかし、寛大さ比率基準は先に検証した連続値の場合と比べて大きな違いを示している。この理由を考察していく。寛大さ比率基準切断プロトコルは $p_i - q_i$ の確率で切断権を獲得できるため、エピソード初期に p 大かつ q 小の進化が促進され、 $S = (0.5, 0)$ が安定な戦略となる。ここで、突然変異によってナッシュ均衡の戦略のエージェント $S = (0, 0)$ が生じた場合を考えると、突然変異エージェントに既存のエージェントが獲得利得で負けてしまう。そのため、 $S = (0, 0)$ の合理的戦略（ナッシュ均衡の戦略）を持つエージェントを排除することができず、その戦略は容易に伝播して最終的にナッシュ均衡に陥ると考えられる。

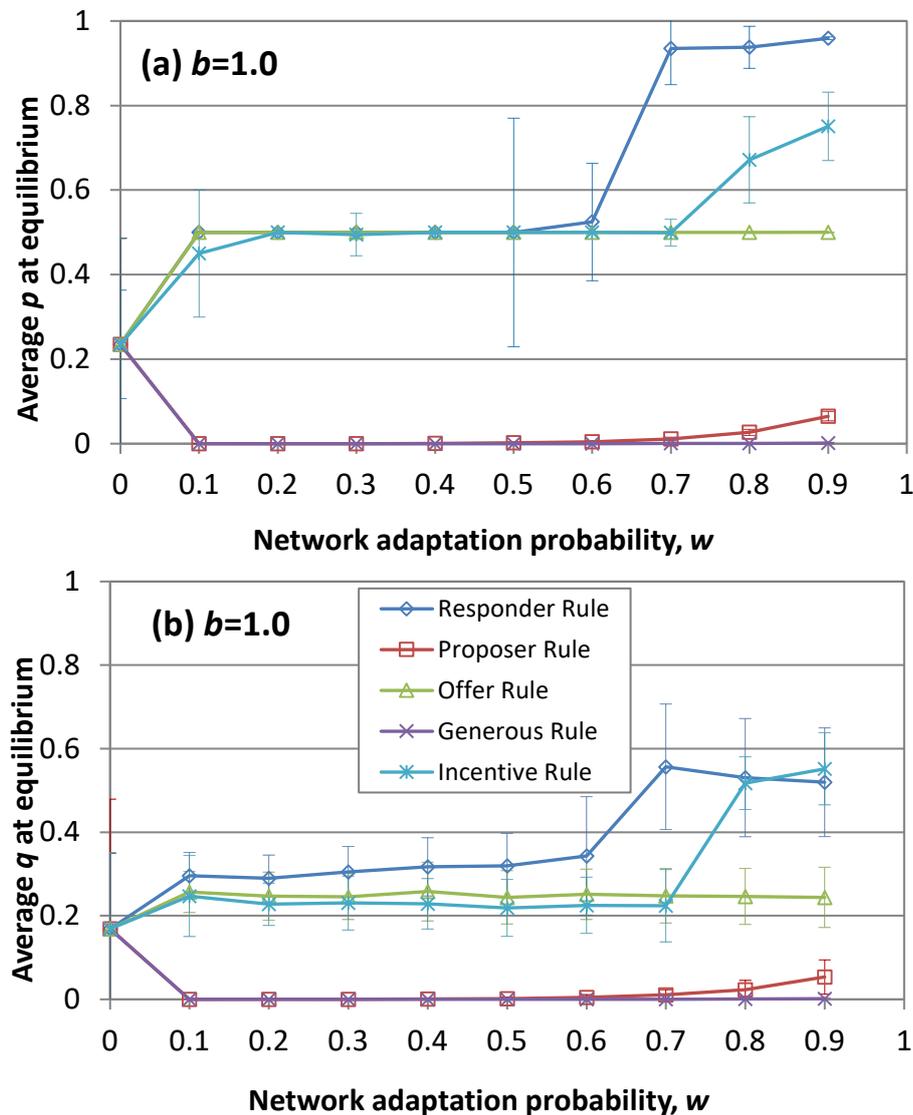


図 3.4.12 それぞれの w における均衡時の(a)社会平均オファーと(b)社会平均許容額.

3.4.4.1.1. 平衡状態

インセンティブ比率基準における $w = 0.8, 0.9$ の場合の結果に注目すると、 $p = 0.7$ 付近で均衡していることがわかる。離散戦略では、戦略値を $p = 0, 0.5, 1$ の離散値でしか定義していないにも関わらず、それ以外の値に均衡するという事は、2つの可能性が考えられる。

1つは、均衡点に **Bi-Stable** 性があることである。結果図は 100 試行平均の値をプロットしているため、例えば、100 試行中約 50 回は $p = 0.5$ の均衡、残りの 50 回は $p = 1$ の均衡に至れば、 $p = 0.75$ 付近の結果となる。

2つ目は、併存平衡である。併存平衡とは、ある 1 試行を見たとき、すべてのエージェントが $p = 0.5$ や $p = 1$ という戦略値を採るようになるのではなく、それらの戦略を持つエージェントが共存することである。この場合、1 試行の社会平均戦略値を見たときでも、中途半端な値となる。

では、このインセンティブ比率基準の均衡状態はどのタイプなのであろうか。結果図 3.4.12 に示されている標準偏差バーを見てみると、そこまで大きく振れていない。このことから、毎試行の均衡値はそこまでばらつきがないと考えられ、均衡点は併存平衡に至っている可能性が高い。また、均衡時のオファー成立率を見てみると、ほぼすべてのオファーが成立していた。このことも併存平衡を裏付ける証拠である。

では一体どのようなトポロジーで均衡しているのか、具体的に見てみる。図 3.4.13 に $w = 0.9$ 、インセンティブ比率基準におけるある 1 試行の均衡時ネットワークトポロジーを示している。この結果図から、同様の戦略値 p を持つエージェント同士がクラスターを形成して併存平衡に至っていることが視覚的に容易に見て取れる。リンクの再接続は、ランダム接続という極めてシンプルなルールであるにも関わらず、このように非常に規則的な均衡状態に至るということは興味深い性質であるといえる。

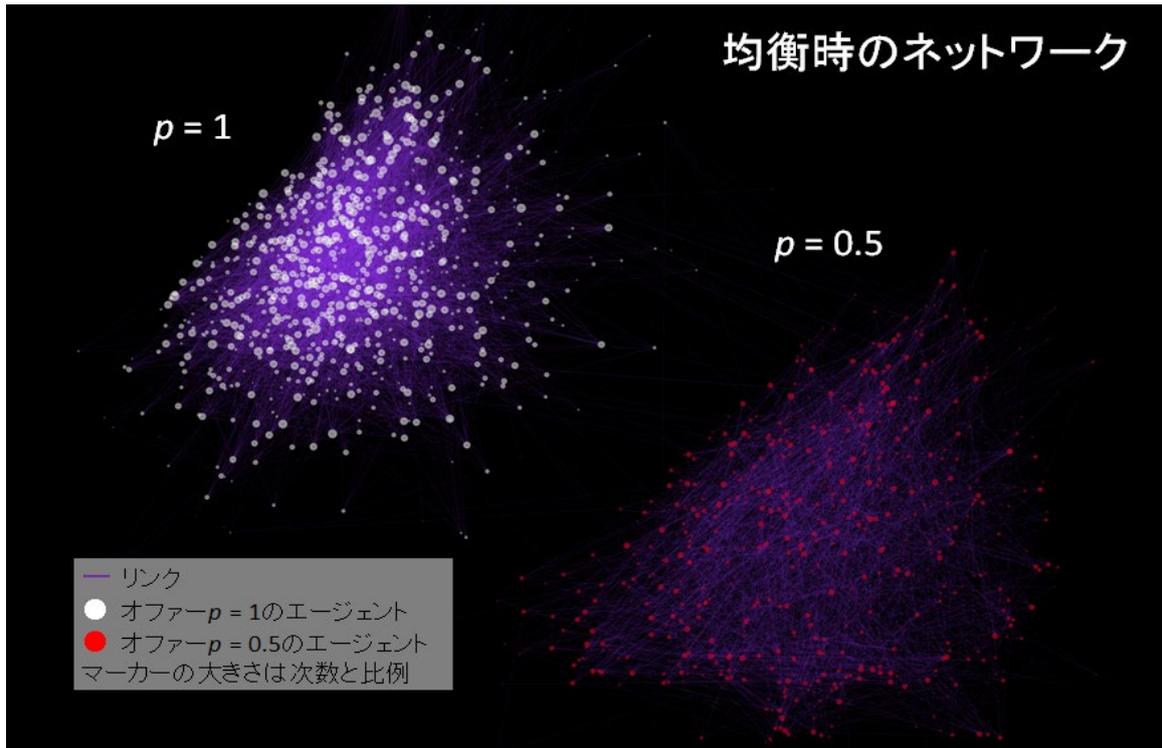


図 3.4.13 $w=0.9$, インセンティブ比率基準切断プロトコルにおけるある 1 試行の均衡時ネットワークトポロジー.

3.4.4.2. $b = 2.0$ の場合

図 3.4.14 に離散戦略, $b=2.0$ のときの社会平均オファーと社会平均許容額を示す. これを見ると, Responder 基準, オファー比率基準, インセンティブ基準は利得構造の影響を受けて p が enhance されていることが見て取れ, 連続戦略のときと同様の傾向を示している. $w = 0.7$ のときの Responder 基準, $w = 0.8$ のときのインセンティブ比率基準を見てみると, 非常に標準偏差の振れが大きく, Bi-Stable な均衡に至っている可能性が非常に高い. 離散戦略定義かつ $b = 1$ のときの結果と合わせ見ると, w が大きくなることで異種の戦略を持つエージェント同士がクラスターを形成して共存できる併存平衡状態になり, その併存平衡に至るか Bi-Stable な均衡に至るかの境目となる臨界ネットワーク適応速度が $w = 0.6 \sim 0.7$ 付近に存在する可能性があるといえる.

Proposer 基準, 寛大さ比率基準は w 大の領域でわずかに $p \neq 0$ となっているが, ほぼナッシュ均衡であるといえる.

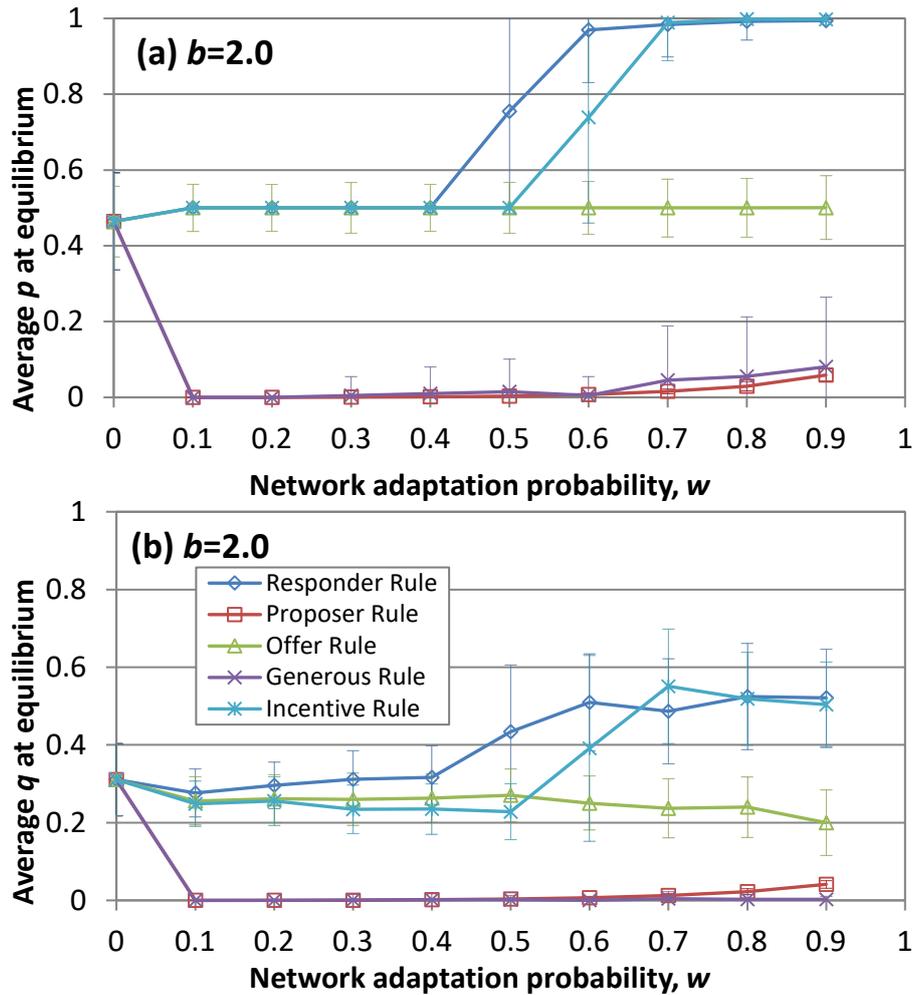


図 3.4.14 それぞれの w における均衡時の(a)社会平均オファーと(b)社会平均許容額

3.4.5. 離散戦略での均衡時におけるネットワークパラメーター

図 3.4.15 に離散戦略, $b = 1.0$ のときの次数相関を示す。

3.4.2 の連続戦略における均衡時のパラメーターと比較すると、極めて異なることがわかる。特にインセンティブ比率基準では、 $w = 0.7$ を超えると、正の相関を示している。一体どういったメカニズムで正の次数相関が創発したのであろうか。

これは 3.4.4.1.1 で述べた併存平衡であることが大きく関係していると考えられる。図 3.4.13 を見ると、 $p = 1$ のクラスターの方が規模が大きい。実際、 $p = 1$ のエージェントと $p = 0.5$ のエージェントの次数を見たとき、 $p = 1$ のエージェントの方が相対的に大きな次数となっていることが確認された。これは、 p 大のエージェントほど切断権を獲得しやすく、次数が容易に増加する切断プロトコルの特性によるものである。さらに、同じ p 同士でクラスターを形成しているため、結果的に次数大のエージェント同士が繋がりあった状態が生まれ、正の次数相関となったと考えられる。

図 3.4.16 に $b=1.0$ のときのクラスター係数, 平均パス長, 島の数といったネットワークパラメーターをまとめて示す. これらのパラメーターは連続戦略のときと比べて特筆すべき変化は生じなかった.

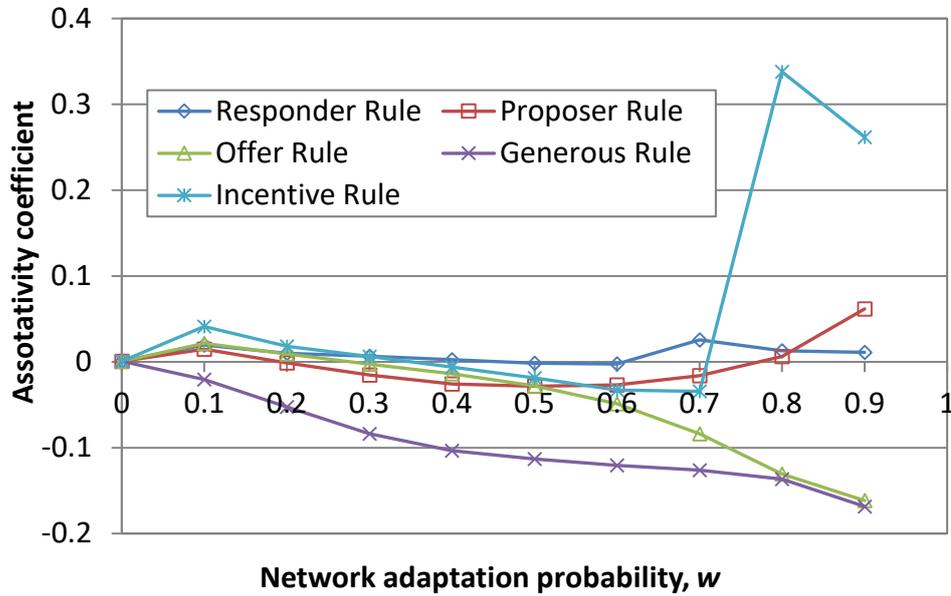


図 3.4.15 各切断プロトコルにおける次数相関とネットワーク適応確率の関係.

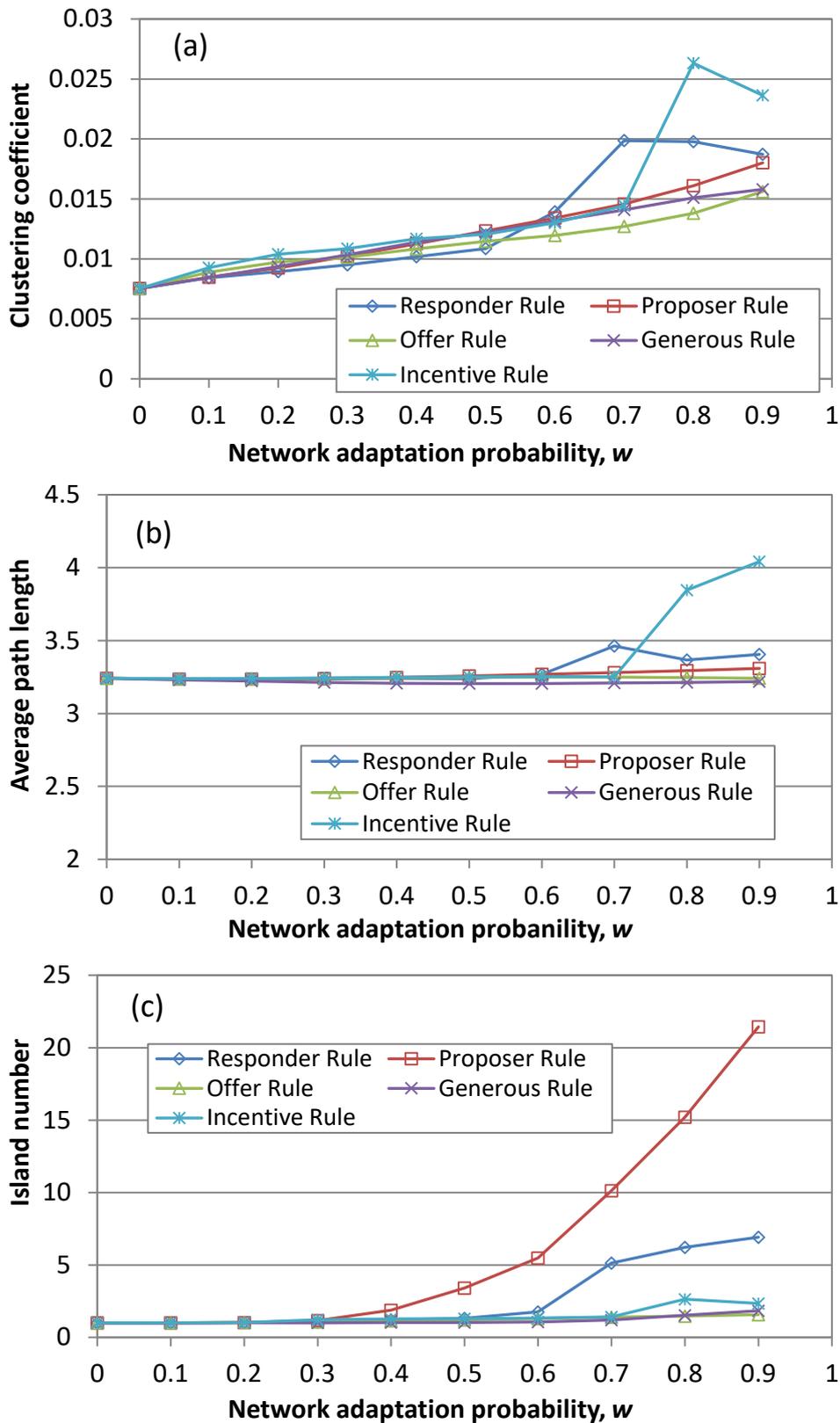


図 3.4.16 各切断プロトコルにおける(a) クラスタ係数(b)平均パス長 (c)島の数とネットワーク適応確率の関係。

3.5. 結論

相手へのオファーが増幅されるメカニズムによりジレンマ性をゲームに付加出来る拡張 **Ultimatum Game** を定義し，実数 2 次元の戦略であるオファーとオファー許容値 $S = (p, q)$ の適応に加え，リンク切断によるネットワーク適応を考慮した共進化モデルを構築した．

ネットワーク適応を決めるリンク切断プロトコルに関して，確定的プロトコル 2 種，確率的プロトコル 3 種を提示して，一連の数値実験により，切断プロトコルによる社会の公平性と効率性の進化への影響を検討した．その結果，切断プロトコルの違いによって，大きく異なる均衡に至ることが明らかとなった．また，モデルを拡張することで切断プロトコルの影響力を計量し，**Responder 基準**や**Proposer 基準**といった確定的切断プロトコルでは，現実世界では起きえない切断イベントを許容することで，過度に気前よい，過度に寛大な均衡に至ることが明らかにした．確率的プロトコルでは，現実にはあり得ない切断の生起は抑制され，被験者実験で観察されるような適度な気前よさ，適度な寛大さを創発させる．

また，他者へのオファー増幅を考慮するとき ($b=2$)，**Responder 基準**が，より効率的な社会均衡へと誘導することが示唆された．しかし，この切断プロトコルは，相手からの切断を免れるためひたすら大きなオファーを出す方向の進化を社会にもたらし，現実的には様々な非合理性を生み，ある種の社会頑強性を消失せしめる（相手に媚びるまでの大きなオファーを指向する）可能性もあり，上記のように現実の人間社会システムでこのようなリンク適応が営まれているとは考えにくい．

離散戦略での実験によって，戦略値と次数の相関が明らかになった．また，この結果を踏まえると，連続戦略の場合でも，離散戦略の様にあからさまな相関は生じないが，非常に細かいタイムスケールで戦略値進化を追ったとき，同様な現象が創発している可能性がある．

切断プロトコルの妥当性を考えたとき，提示中のインセンティブ基準は，デフォルトモデルでも十分に公平なオファーを供出し，かつ **Responder 基準**に次ぐ高い社会効率性を達成出来る上，自他のオファーと許容閾値の大小関係に応じてリンク切断権付与する，社会的には「理に適った」プロトコルであると言えよう．

4. 記憶に基づいた確率的意思決定が協調行動に与える

影響

4.1. 研究背景と着眼点

Miyaji ら[102]が報告した内容を中心に、ネットワーク互惠下において、協調の進化とエージェントの保有する記憶長の関係を詳細に検証する。現実社会では、人と人との相互作用や状況に応じて科される社会ジレンマは変化に富むものでありながら、人間社会システムには頑強な協調が維持されている。このことから、如何なる状況下でも高度な協調を維持できる普遍的な協調創発機構が背後に機能していることが想起される。ここにおいて、囚人のジレンマゲーム (PD ; Prisoner's Dilemma Game) やチキンゲーム (CHG ; Chicken Game) を単独に切り出し、ネットワーク互惠や記憶長戦略を別々に検討している従来のアプローチとは異なる統合的視座が要請される。本章では、ネットワーク互惠と直接互惠を組み合わせ、さらに不確定状況に柔軟に応じることを担保する確率性を加味した空間型 1 記憶長混合戦略ゲームを提示し、PD, CHG といった異なるジレンマクラスに応じて協調行動は創発され得るのかを検証する。

4.2. 研究概要

通常進化ゲーム理論研究においては、ネットワークで規定された隣人に対して協調または裏切りのどちらかを選択するバイナリーなモデルが一般的である。本章では、エージェントに記憶の概念を付与し、相手の前手に応じて自分の次手を確率的に決定する一記憶長モデルを空間型 2×2 ゲームに組み込んだモデルを提示する。囚人のジレンマゲームでは、互いに C を出す R 互惠が、チキンゲームでは C と D をある割合で出し合う ST 互惠が社会利得を最大化する上で最も優れた互惠機構であるが、提示モデルによるシミュレーションから、エージェントはそれぞれのジレンマクラスに応じて社会利得を最大化するために最適な戦略へと進化していく事を報告する。

4.3. モデル

4.3.1. 2×2 ゲーム

2 人 2 戦略 (2×2) ゲームを用いる。本モデルでは、2 実数値からなる戦略 $s_i = (p_i, q_i) \in [0,1]$

を導入する. p_i は, 対戦相手が前回 C だった時, 次回協調する確率, q_i は, 相手が前回 D だった時, 次回協調する確率である. 戦略としては, p_i や q_i のような確率ではなく, バイナリーに離散値 0 (D) または 1 (C) で手を直接規定する考え方が一般的であるが, 本章では人のランダムな振る舞いをモデル化するため, 確率的な意思決定方法として実数値戦略を用いる [104]. このモデル設定では, エージェントは対戦相手の前回の振る舞いに応じて異なる戦略値を参照して次手を決定するため, 対戦相手の前回の手を 1 回分記憶することになり, C か D かの手 (行動) は p_i と q_i によって確率的に決まる. すなわち 1 記憶長混合戦略である. エージェントは, 自分と相手の振る舞いのに応じて利得を獲得する. もし, 自分と対戦相手共に C 戦略を選択した場合, Reward (R) を, 両者共に D を選択した場合, Punishment (P) をそれぞれ利得として得る. もし, 一方が C, 他方が D 戦略を用いた場合, 前者は Sucker (S), 後者は (T) を得る. これら利得 R, S, T, P の大小関係を変化させることで, 様々なジレンマ構造をモデル中に再現できる. Tanimoto & Sagara [103] のスケーリング方法に倣い, $R=1$, $P=0$ と固定することで,

$$\begin{pmatrix} R & S \\ T & P \end{pmatrix} = \begin{pmatrix} 1 & -D_r \\ 1 + D_g & 0 \end{pmatrix} \quad (4.1)$$

と表すことが出来る. これによってジレンマ構造を D_g (チキン型ジレンマの強さを意味する) D_r の (Stag-Hunt 方ジレンマの強さを意味する) の 2 変数のみによって操作できる. この時, PD ゲームにおける利得構造は $T > R > P > S$ すなわち, $D_g > 0$, $D_r > 0$ の領域, CH ゲームにおける利得構造は $T > R > S > P$ または $T > S > R > P$ すなわち $D_g > 0$, $D_r < 0$ の領域で表すことが可能となる (図 1 を参照). 社会粘性がない状況下で R 互惠に比して ST 互惠が有意となるのは CH でも $S + T > 2R$ の条件を満たす場合である.

4.3.2. シミュレーション手順

毎シミュレーションエピソード初期, 各エージェントをネットワーク上に配置し, $[0, 1]$ の範囲でランダムに戦略 $s_i = (p_i, q_i)$ を付与する. シミュレーションの第 1 ステップ目の初手として, C または D をランダムに割り当てておく. ゲーム及び戦略適応の進行は以下に述べるアシンクロ更新に基づく. すなわち, 初期設定が完了した後, ランダムに選択されたエージェント i が全隣人と対戦することで積算利得 π_i を得る. 各ゲームでは戦略に応じ, 各相手の前手により C か D かを確率的に決める. 次に, ランダムに選択した隣人 j の積算利得も同様に求めておく. 戦略適応ではエージェント i は, 隣人 j の利得 π_j と自利得 π_i とにより, 自戦略を保

持するか j の戦略 $s_j=(p_j, q_j)$ をコピーするかを以下に基づき確率的に決める (Pairwise Fermi 更新; 例えば[105]). エージェント j の利得が i よりも高ければ高いほど j の戦略をコピーする確率が上昇する.

$$P_{i \leftarrow j} = \frac{1}{1 + \exp\left[\frac{\pi_i - \pi_j}{\kappa}\right]}, \quad (4.2)$$

ここで, κ は温度係数と呼ばれ, 値が大きいほど利得差の影響が強くなり, コピーの不確実性が低下する. 本章では, $\kappa = 0.5$ とする.

以上の過程を均衡に至るまで繰り返し, 均衡前 100 ステップの社会平均 p , 社会平均 q , 協調率 ρ_c (C の実際の出現頻度) を考察する. 本章では, 次数 $k=4$, $N=70 \times 70 = 4900$ の周期境界格子グラフ, または次数 $k=4 \sim 24$, $N=4900$ のランダムレギュラーグラフを基盤ネットワークとして用い, 均衡時の戦略値と協調率を 100 試行アンサンブル平均したものを結果とする.

4.4. 結果と考察

はじめに, 本シミュレーションにおける結果を俯瞰するため, 図 4.4.1 に示す全領域の均衡時の協調率と行動プロファイル中の戦略値 p, q を見てみる. 協調率を見ると, レプリケータダイナミクスによって導出される解析解の結果と比べて著しく協調が **enhance** されている. 従来の 1bit 戦略表記によるネットワーク互惠のモデルでは $D_g = D_r = 1$ の PD では裏切り相に吸引されるが, 本モデルではほぼ完全な協調相が達成されている. 戦略値 p と q を見ると, CHG の一部領域を除いて p は全域で 1 に近い値となっており, q はジレンマ強さによって異なるが, ジレンマが強い $D_g > 0$ の領域では 0 に近い値に均衡している (Fig.4.4.1(c)). つまり, D_g と D_r が大きくなってジレンマが強くなり, 裏切って T をとるインセンティブが大きくなると, 裏切者に対しては裏切り返す戦略進化によって R 互惠を維持していることが分かる. 本モデルでの戦略値進化の帰結は Tit-for-Tat 戦略(TFT) [15] に似ているが, 裏切りに対して常に裏切り返すわけではないという戦略となっており, TFT の弱点である一人の裏切者によって全員が裏切り相に吸引されてしまう欠点を克服したものとなっている. このような戦略へと均衡することが出来るのは, 実数値 p と q により C or D に関して所謂混合戦略[58]となっているためである. 図は示していないが, 1 記憶長離散戦略 $s_i = (p_i, q_i) \in \{(0,0), (0,1), (1,0), (1,1)\}$ を用いた結果との比較から, 戦略決定プロセスの確率性が, 頑強な協調 **enhance** 効果に寄与していると云える.

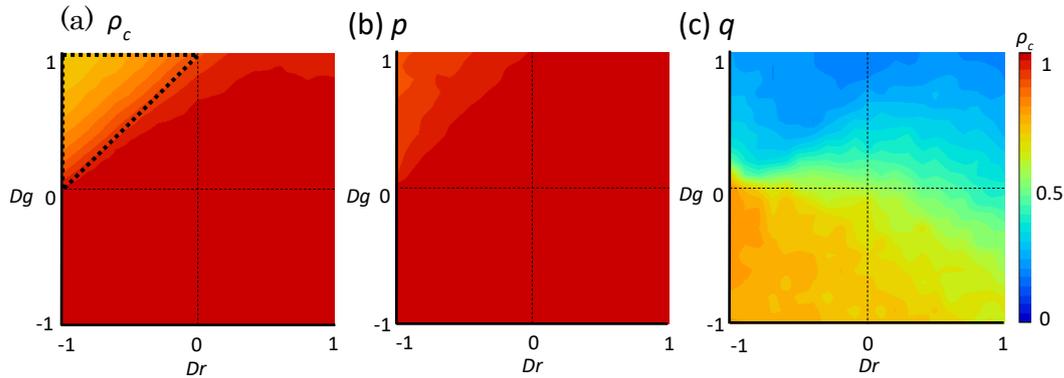


図 4.4.1 均衡時の協調率と戦略値 p , q . 縦軸は D_g , 横軸は D_r を表している. 各プロットは 100 アンサンブル平均の値を示している. 本シミュレーションではプレイヤーは隣人に対して異なる戦略を用いることが出来るため, 協調率の導出をする際はフォーカル 1 人に対してネットワークで規定された隣人 4 人の中から 1 人ランダムに選択し, その隣人に対してどの戦略を採ったかを観察し, その平均値を 1 ステップの協調率としている. 右上のセクションは Prisoner's dilemma game (PDG), 左上が Chicken game (CHG), 右下が Stag-hunt game (SHG), 左下が Trivial game (TG) と呼ばれるゲーム構造を有している. CHG で, 黒点線で囲まれている領域は $2R < S + T$ を満たす為, 完全な協調状態よりも中途半端な協調率を維持することでより高利得を上げることが出来る ST 互惠や alternate reciprocity と呼ばれる現象が創発する.

CHG の結果に注目してみると, 記憶長があるエージェントを以てしても CHG の一部分 (図 1 中の点線で囲まれた部分) は協調率 1 に至っていない. この領域ではいったい何が起きているのかを詳しく見ていく.

ST 互惠が R 互惠よりも優位となるのは, well-mixed を前提として考えたとき, $2R < S + T$ を満たす領域, すなわち図 4.4.1 中に黒点線で示されている $D_g > D_r + 1$ の領域であり, 本モデルで唯一協調率が 1 付近に到達していない領域と一致する. 従って, 本モデルでも何か特別な現象が創発している可能性があり, この結果について証左に検討する.

R 互惠よりも ST 互惠を築くメリットを明確に把握するために, ST 互惠が社会の利得を最大化する上で有意となる領域では, エージェントの平均利得と協調率にどのような関係があるのかを平均場近似を用いて導出する. 均衡時の協調率を ρ_c としたとき, ネットワーク中のある 1 プレイヤーの期待利得を考える. あるプレイヤー組が R をとる確率は ρ_c^2 , S をとる確率

は $\rho_c(1 - \rho_c)$, T をとる確率は $(1 - \rho_c)\rho_c$, P をとる確率は $(1 - \rho_c)^2$ となるから, あるプレイヤーが隣人1人と対戦した時に獲得する期待利得 $\langle \pi \rangle$ は,

$$\begin{aligned} \langle \pi \rangle &= \rho_c^2 R + \rho_c(1 - \rho_c)S + (1 - \rho_c)\rho_c T + (1 - \rho_c)^2 P \\ &= (R - S - T + P)\rho_c^2 + (S + T - 2P)\rho_c + P \\ &= (R - S - T + P) \left\{ \rho_c + \frac{S+T-2P}{2(R-S-T+P)} \right\}^2 - \frac{(S+T-2P)^2}{4(R-S-T+P)} + P \quad (4.3) \end{aligned}$$

となる. 従って, 最大利得を獲得できる協調率と最大利得は,

$$\rho_{c_{max}} = \frac{S+T-2P}{2(R-S-T+P)} \quad (4.4)$$

$$\langle \pi \rangle_{max} = \frac{(S+T-2P)^2}{4(R-S-T+P)} + P \quad (4.5)$$

となる.

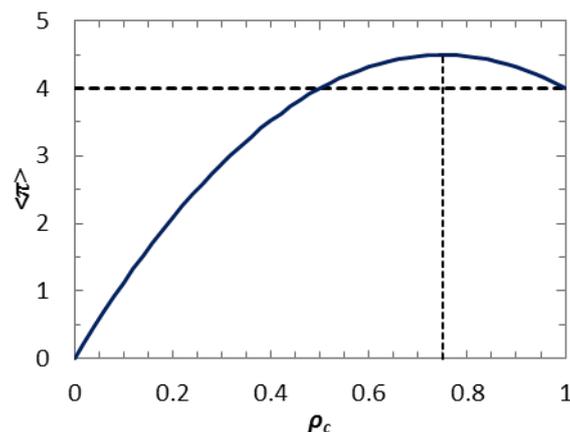


図 4.4.2 $D_g = -D_r = 1$ のときの各協調率における利得曲線. Well-mixed を仮定して平均場近似することにより導出され, 利得は協調率に対する 2 次関数となることが分かる. 従って, $2R < (S + T)$ を満たす領域ではジレンマ強さごとに社会利得が最大となるような協調率が $0 < \rho_c < 1$ の範囲に存在する. 例えばこの図が表すように, $D_g = -D_r = 1$ のとき, 完全な協調状態である $4R$ のときに生み出される社会利得を超えた最大社会利得 $E_{max} = 4.5$ が $\rho_c = 0.75$ (点線) のときに生み出されることがわかる.

図 4.4.2 は $D_g = -D_r = 1$ のときの $\rho_c - \langle \pi \rangle$ 関係である. このように空間構造を考えない平均

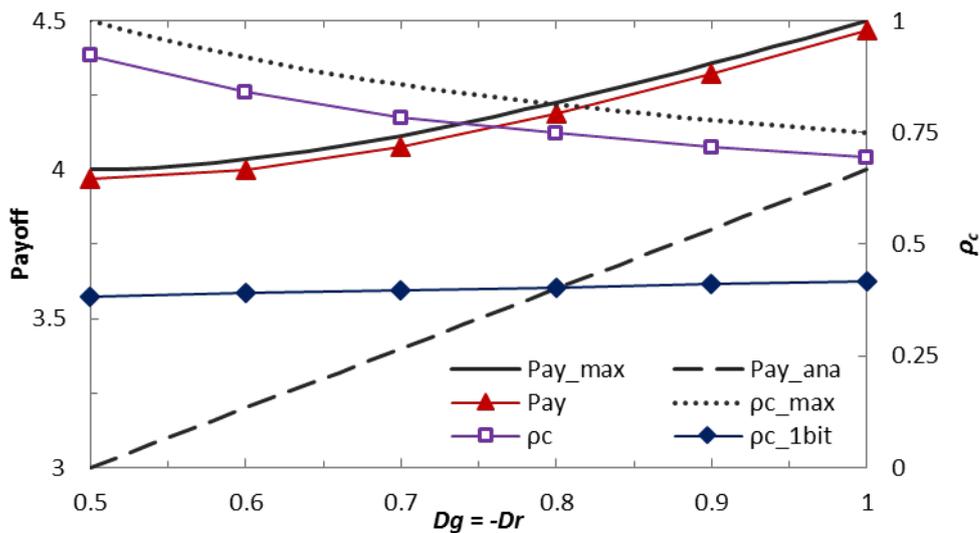


図 4.4.3 CHG における均衡時の平均利得と協調率を平均場近似で導出した理論最大利得と協調率と比較したもの. 黒実線が平均場近似によって導出された最大利得の軌跡曲線で, 黒点線がその時の協調率の軌跡曲線. 黒破線はレプリケータダイナミクスで予測される well-mixed のときの平均利得. \triangle と \square は本シミュレーションの結果である.

場近似の元では, 平均利得は協調率の 2 乗に比例するため, 社会利得を最大化する協調率が $0 < \rho_{c_{max}}$ の範囲に出現する. また, 空間構造があつたとしても, 実数値戦略により, 均衡時にはほぼ一様な空間分布になる場合には, ここで示した平均場近似が成立するものと期待される. この(4.4), (4.5)式で示される理論上の最大利得とジレンマ強さとの関係を $D_g = -D_r = x$ として表すと以下となる.

$$\begin{aligned} \rho_{c_{max}} &= \frac{S + T - 2P}{2(R - S - T + P)} \\ &= \frac{2x + 1}{4x} \end{aligned} \tag{4.6}$$

$$\begin{aligned} \langle \pi \rangle_{max} &= \frac{(S + T - 2P)^2}{4(R - S - T + P)} + P \\ &= \frac{(2x + 1)^2}{32x} \end{aligned} \tag{4.7}$$

但し、 $R < (S+T)/2$ の時のみこの式は有効であることに注意する。何故なら、 $R \geq (S+T)/2$ では、 ST 互恵よりも互いに C を出す R 互恵が常に最大利得を生み出すからである。

ST 互恵における理論上の最大利得と本モデルによる利得とを比較したのが図 4.3.3 である。図では協調率と平均利得を示している。 **well-mixed** を仮定したレプリケータダイナミクスによって導かれる解析解における協調率は常に 0.5 となる。本モデルでは、平均利得はほぼ理論上の最大利得に近い。PDG と異なり、CHG では空間構造があることが C -cluster の成長を阻害する可能性があることが知られている[106]。このことから類推すると、空間構造は ST 互恵を達成する上でも阻害要因になる可能性があるかもしれない。しかし、本モデルのように、空間構造に加え、エージェントに 1 記憶長の混合戦略を許容して、隣人の前手から次手を確認的に決定する機構を付与してやれば、ほぼ理論上最大の効率で ST 互恵を達成できることがわかった。

本モデルでは、前記したように空間構造によるネットワーク互恵と記憶長戦略による直接互恵が組み合わさって機能している。図 4.4.4 は、本モデルの基盤グラフのトポロジーを大次数の **Random Regular Graph (RRG)** とすることで、擬似的に記憶長戦略による寄与だけをみたものである。縦軸は協調率と平均利得を表し、比較のためレプリケータダイナミクスによる均衡協調率（平均利得）と 1bit 戦略による通常のネットワーク互恵によるそれらとを併記している。また、 $D_g = -D_r < 0.5$ では R 互恵、すなわち協調率 1 が最大利得となる点を留意する。次数が大きくなり空間構造の影響が小さくなると、 R 互恵が望ましい領域でも協調率 1 を維持し得なくなり、 ST 互恵が望ましい領域では、理論上最大利得となる協調率を大きく下回り、1bit 戦略による協調率をも下回る。以上のことから、本モデルにより、CHG において R 互恵と ST 互恵が効率的に達成されるためには、記憶長戦略だけでなく空間構造の存在が重要であることがわかった。

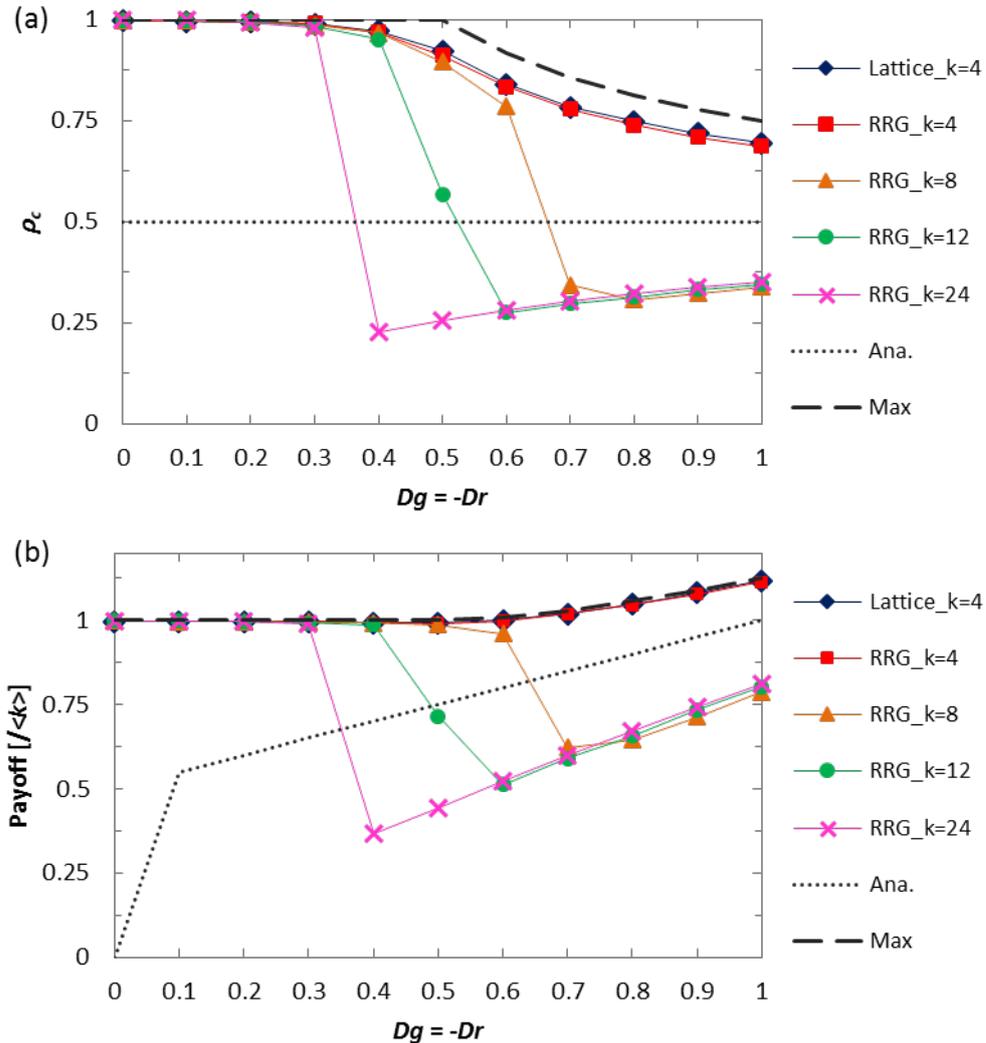


図 4.4.4 Random regular graph (RRG) を用いたときの(a)協調率と(b)平均利得. ネットワークが仮定されている場合の平均利得は, 次数によって規格化された, 1 リンク当たりの平均利得を表している. 点線は well-mixed を仮定した時平均場近似で予測される最大利得と, レプリケータダイナミクスによって予測される解析解. RRG グラフの平均次数を大きくしていくと, ネットワーク互惠の効果が減少し, well-mixed の状態へと近づいていくと云え, その傾向を読み取ることでメモリーの効果のみを検証することが出来る.

次に PDG に注目する．このジレンマクラスでは，CHG のように S と T を取り合うことが社会利得を最大化するわけではなく，純粋に協調を維持することが双方にとって最も利益をもたらす公平なパレートである．PD 領域における通常の 2×2 ゲームでは，ジレンマが大きくなると T と S の差が大きくなり，ある程度の社会粘性があったとしても R 互惠を維持できなくなって全員が裏切り状態に吸引されてしまう．一方，本モデルによる結果を見てみると， $0 \leq D_g, D_r \leq 1$ の PDG 全域でほぼ協調率は 1 となっており，公平なパレート最適を達成している．

本枠組みではどのようにしてジレンマが強い領域で協調を維持しているのだろうか．図 4.4.5 に示されている $D_g = D_r = 1$ における Snapshot を見てみる．図から，初期に p, q 共に小さな値である D 戦略型のプレイヤーが増加しており，徐々に全員裏切り相へと引き寄せられていっていることが分かる．しかし，10Step を過ぎた辺りから徐々に p が比較的大きめのエージェント同士でクラスターを形成し， R を取り合って利得を維持することで勢力を拡大していく．そして最終的にはほとんどのエージェントが C を出すようになる．

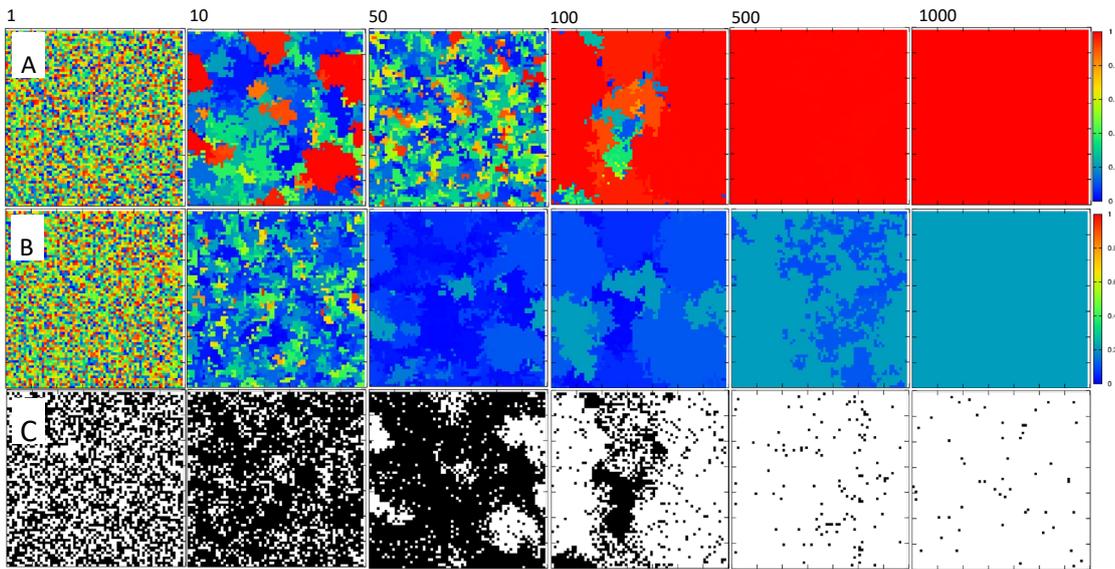


図 4.4.5 $D_g = D_r = 1$ (PDG)の時のスナップショット．図 A が戦略値 p ，図 B が戦略値 q である．図 C は対象ステップに各プレイヤーがどの戦略を採ったのかを表しており，白が協調戦略，黒が裏切り戦略を採ったことを意味する．フォーカル 1 人に対してネットワークで規定された隣人は 4 人いるが，今回はフォーカル 1 人に対してその隣人を 1 人ランダムに選択し，その隣人に対してどの戦略を採ったかを表示している．

この Snapshot からわかるように，本モデルでは，プレイヤーは自分の持っている混合戦略

値に基づき相手の前手から自身の次手を決定するため、 p が 1、 q が 0 ように戦略が完全に一致するエージェント同士でクラスターを形成しなくても、比較的 p が大きめのエージェント同士でまとめれば C を出し合うクラスターを形成することができる。つまり、確率的な戦略表記によって各エージェントは多様な行動が許容され、C を出し合うクラスターが形成できる条件を緩和しているといえよう。さらに、戦略値として p と q の 2 パラメーターを保持しているため、協調者に対する戦略と裏切者に対する戦略を使い分けることが大きな advantage として機能する。すなわち、 p 大のクラスターを形成して拡大していく際に、クラスターの境界にいるエージェントはクラスターの内側の協調者には C を出して R を取り合う一方で、クラスターの外側の裏切者に対しては D を出すことで Punishment を加え、彼らに大きな利得をとらせないことで、この協調型クラスターを頑強にしている。PD 領域で協調型クラスターが安定に存在できない理由は、協調者が裏切者と対戦した時に、利得を獲得できない可能性があるからであるが、本モデルのように記憶長があれば、裏切り者か協調者か対戦相手の前手に応じた自手の使い分けによって裏切られたときのコストを低減させることができる。これにより、ジレンマが強い領域でも C 相の維持が可能となる。ここで、協調的戦略が安定に存在するためには、 q の値が 0 ではないことが非常に重要である。本モデルでは確率的に手が決定されるため、 p が大きめの値の戦略値を保持しているプレイヤーでも p が完全に 1 でない限りまれに協調者に対して裏切り戦略を採る場合がある。その時に、相手の q が 0 であると裏切りが裏切りの連鎖を呼ぶことになる。図 4.4.1 を見ると、 $0 \leq D_g, D_r \leq 1$ の領域では q は 0.2 – 0.45 付近に均衡しており、戦略は、相手の裏切りに対して毎回裏切り返して報復を加えるのではなく、相手を罰することを若干躊躇う戦略に進化していると云える。この進化の帰結は TFT が内包していた、一人の裏切者がトリガーとなって全員が裏切り戦略に陥ってしまうと云った弱点を克服するものとなっている。図 4.4.6 を見ると分かるように、さらにジレンマが強くなると ($D_g=D_r>1$)、S をとらされた時の利得減少を防ぐために q の値が小さくなる方向へ、すなわち非寛容な戦略へと進化していくため、次第に協調率は減少する。とはいっても、非常にジレンマの強い $D_g = D_r = 5$ でも協調率は 0.5 程度を維持している。

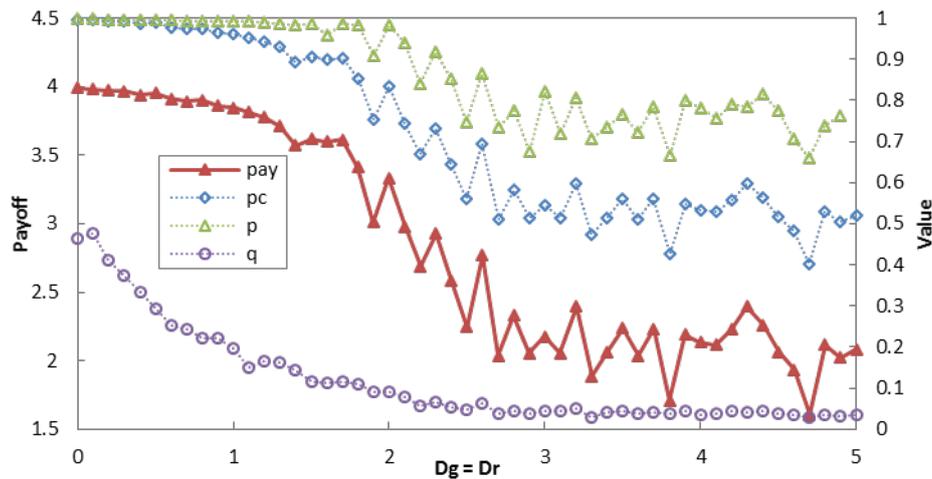


図 4.4.6 囚人のジレンマゲームにおける均衡時の平均利得 (“pay”), 協調率 (“pc”), p , q . $D_g = D_r = 1$ を超えた辺りから, p が徐々に 1 から低下していく. それに伴って協調率と平均利得が大きく減少し始める. しかしながら $D_g = D_r = 5$ のような極めてジレンマの強い領域においてもある程度の協調関係が維持できている.

4.5. 結論

Square Lattice 上の各プレイヤーに 2 実数値で表現される 1 記憶長を許容する進化ゲームモデルを提示した. エージェントは自分の置かれた状況 (ジレンマクラス) に応じて出す手を変え, 非常に大きなジレンマ下でも高い社会平均利得を達成する互惠機構が創発することが示された.

C と D を適当に出し合って ST 互惠を取ることが望ましいチキンゲームの領域では, 本モデルにより平均場近似から導かれる理論上の最大値にほぼ等しい高い社会利得を達成できる. この頑強な互惠機構は, 空間構造によるネットワーク互惠と記憶長戦略による直接互惠が一体的となって効果を発揮することからもたらされる.

常に C を出し合って R 互惠を取ることが望ましい PD ゲームの領域では, 従来のネットワーク互惠モデルで検討されてきたジレンマより遙かに強い状況下でも, 本モデルにより協調が維持される. これは 1 記憶長戦略により, 協調的クラスターの境界にいる協調者が, 対戦相手が協調者か裏切り者かに応じて手を出し分けられることが大きく寄与している. また, 実数値で定義された行動プロファイルの進化の帰結を見てみると, 直接互惠戦略として代表的な TFT がノイズに対して脆弱である欠点を克服したような行動様式となった. Tanimoto & Sagara [103]の結果から推量すると理想的な ST 互惠には 2 記憶長戦略による直接互惠が必要であると思われるが, 本モデルのように, 空間構造を導入し, 実数値表記により確

率的な戦略が定義できるのならば、1 記憶長戦略で十分であることがわかった。

この結果は条件設定の影響をどれだけ受けているかを検証するため、フェルミ関数の温度係数を調節することで戦略更新時のランダムネスを変化させて同様のシミュレーションを行ったが、最終的な結果には大きな違いは現れなかった。さらに、コピーノイズの影響を検証した結果、この影響はほとんどなかった。従って、本実験においてはノイズの有無や戦略更新に内包されるランダム性よりも前手に基づいて確率的に戦略を決定することが出来ることに依る各プレイヤーの行動の多様性の方が社会利得を最大化するための協調的振る舞いをサポートする上で遥かに大きな影響を持っているといえる。

5. 混合戦略モデルにおける戦略更新方法に内包される

確率的揺らぎと協調創発の関係

5.1. 研究背景と着眼点

進化ゲーム理論に基づいたシミュレーションモデルを構築し、囚人のジレンマゲームにおいてどのような効果が協調創発に寄与しているのかを検証する。本章では Miyaji ら[107]が報告した内容を中心に、混合戦略進化ゲームモデルにおける更新方法の設定が系内の協調進化の様子に与える影響を検証する。現行、混合戦略ダイナミクスにおける更新プロセスでは、完全な同時性を考慮したものが多数である。離散戦略定義では、同時性の有無が帰結に大きな影響を与えることが分かっており、混合戦略定義でも何か特筆すべき影響が生じる可能性がある。本章では、ゲームダイナミクスにおける同時性は、混合戦略定義においてどのような影響を及ぼすのかを検証する。

5.2. 研究概要

プレイヤーに連続手を許容した連続戦略 2×2 ゲームや、メモリーを仮定した 2×2 ゲームでは、プレイヤーは多様な行動が可能となり、それによって協調がエンハンスされることが分かっている。本章では、こういった多様な行動を可能にするゲームモデルの一つ、混合戦略 2×2 ゲームについて詳しく検討する。混合戦略定義の 2×2 ゲームにおいては、従来の戦略適応を行う更新プロセスに加え、各自の保有する戦略値に基づいて行動 (offer, action, behavior) を更新するプロセスがある。戦略適応のタイミングを規定する更新方法は `synchronous update` と `asynchronous update` の二種類あり、混合戦略ゲームにおける行動更新プロセスに関して、同様に規定する必要がある。本章では、この行動更新方法が `synchronous` か `asynchronous` かという事に起因する確率的ゆらぎ (stochastic perturbation) が最終的な協調の進化に大きく影響することを明らかにした。特に、`synchronous` に行動更新が行われる場合、平均次数が大きくなるほど協調的社会へと漸近することを発見した。

5.3. モデル

5.3.1. 2×2 ゲーム

2×2 ゲームを基盤モデルとして用いる。混合戦略定義ではエージェント i は 0 から 1 の間の

実数値で定義された戦略 s_i を保有する．本章では s_i はエージェント i が協調行動 (C) をとる確率を表す．従って， $s_i=1$ のとき，エージェント i は常に協調する協調者，一方 $s_i=0$ のとき常に裏切る裏切者を表す． $D_g=T-R$ ， $D_r=P-S$ とし，簡単のため $R=1$ かつ $P=0$ とすることで利得表は

$$\begin{pmatrix} R & S \\ T & P \end{pmatrix} = \begin{pmatrix} 1 & -D_r \\ 1+D_g & 0 \end{pmatrix} \quad (5.1)$$

と書き表せる．本章では $T>R>P>S \wedge 2R>T+S$ を満たす $0 \leq D_g \leq 1$ かつ $0 \leq D_r \leq 1$ である囚人のジレンマゲームを仮定する．

5.3.2. ネットワーク

エージェントがゲーム対戦を行う相手を規定する基盤ネットワークとして次数 k の周期境界条件のレギュラーラティスネットワーク（二次元正方格子グラフ）を用いる．

5.3.3. 戦略更新と行動更新

混合戦略ゲームの場合，ネットワークで規定された隣人とゲームを行う際，各エージェントは自身の所有する戦略(値)に応じて自身の次手を決定する必要がある．手を決定するタイミングはシンクロ的行動更新 (synchronous process for updating action) と，アシンクロ的行動更新 (asynchronous process for updating action) の二通り考えられる．シンクロ的行動更新の場合，ある一ステップで全エージェントが一斉に自身の持つ戦略 s_i に応じた確率に基づき手を更新し，その手を用いて全隣人とのゲームを行い，利得を計算する．アシンクロ的行動更新の場合，同一ステップ中に，フォーカルは隣人ごとに自身の手を自身の持つ戦略値に基づき更新する．つまり，シンクロ的行動更新の場合は行動 ; C または D を出すかの更新がステップ毎に行われるのに対して，アシンクロ的行動更新の場合，それよりもさらに細かい間隔，すなわち隣人毎に行動が更新される．

戦略プロファイル s_i 自体を更新する戦略更新にもシンクロ更新とアシンクロ更新がある．エージェントの利得計算方法は，戦略更新がシンクロ更新とアシンクロ更新かによって異なる．シンクロ更新の場合，ゲームは以下のように繰り返し実行される．まずエージェント i が全隣人と対戦し利得 π_i を獲得する．同様にして全エージェントがそれぞれの隣人とゲームを行い利得が積算される．このとき，各エージェントがとる実際の行動 C または D は上述の手順で更新されるとする．最後に全エージェントが同時に利得に基づいて戦略 s を更新する．

アシンクロ更新の場合、集団中からランダムにフォーカルエージェント i を選択する。フォーカル i とその全隣人の利得を計算し、積算利得に基づいて戦略適応を行う。同様にして次のフォーカル i をランダム選択し、この過程を全エージェントに対して繰り返す。

本章では戦略更新方法として **Imitation Max (IM)** を用いる。

5.3.4. シミュレーション設定

各シミュレーションは以下のように勧められる。まず各エージェントに戦略 s_i がランダムに付与される。C または D をどのように更新するかは上述のいずれかのプロセスによる。初期設定後均衡に至るまでシミュレーションを継続し、均衡前 100 ステップの平均協調率を均衡協調率として得る。混合戦略ゲームでは確率的に C または D が決定されるため同一ステップ中であって行動を変更する可能性がある。そこで今回は各ステップにおける協調率を近似的に得るため、エージェントが隣人集団中のある隣人 j にとつた実際の行動の平均を協調率として得る。統計的有意性を確保するため、各ジレンマ強さ 100 試行平均した均衡協調率を示す。

5.4. 結果と考察

初めに混合戦略定義は離散戦略定義と比べてどういった違いがあるのかを俯瞰する。図 5.4.1 に離散戦略 ($s_i = 0$ or 1 のエージェントのみ許容する場合)、混合戦略それぞれの戦略定義を用いたときの協調率を示す。次数は $k=8$ である。なお、混合戦略では **asynchronous process for updating offer** によっている。混合戦略を用いた場合、離散戦略定義の場合よりも協調が **enhance** されることが分かる。混合戦略定義の場合は多様な戦略値に基づいて確率的に自身の行動を決定するため、あるエピソードにおいてあるエージェントが一貫した戦略を採り続けるとは限らない。通常の離散戦略定義ゲームでは、エージェントが獲りうる利得は、C エージェントの場合 R か S 、D エージェントの場合 T か P の二通りであるのに対し、混合戦略定義では **cooperator** であっても **defector** であっても P, R, S, T すべての利得を獲る可能性がある。この利得の多様性によって C 型エージェントが一方向的に D 型エージェントとの繋がりによって S をとられ、利得が減少するという状況は回避される。この効果により、離散戦略定義の場合よりも協調がエンハンスされる。

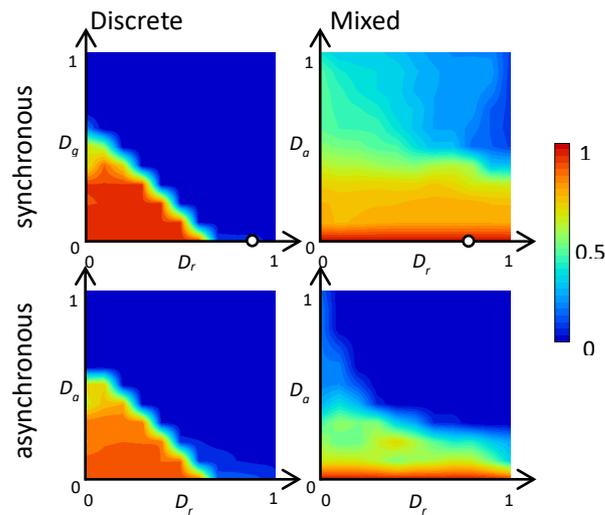


図 5.4.1 離散戦略定義と混合戦略定義における均衡協調率の比較. 基盤ネットワークは次数 8 のラティスである. 上二つのグラフは戦略更新プロセスがシンクロの時, 下の二つは戦略更新がアシンクロに行われたときの結果を表している. 混合戦略における行動更新ルールはアシンクロ更新を用いている.

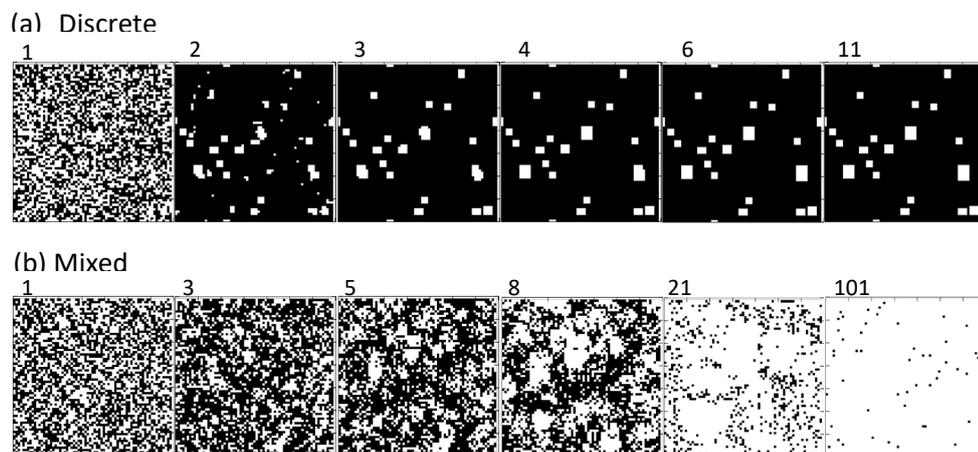


図 5.4.2 格子グラフ上のエージェントが実際に用いた行動の時系列遷移. 協調行動を取ったエージェントを白, 裏切り行動を取ったエージェントを黒で表している. ジレンマ強さは $D_g=0$, $D_r=0.8$ である. パネル(a)は離散戦略定義, (b)は混合戦略定義の時のスナップショットである. 混合戦略定義における行動更新ルールはアシンクロ更新を用いている.

これを確認するために図 5.4.2 に示すスナップショットを見てみる。Pure な C か D のみが許容される離散戦略定義の場合、ジレンマが強い場合は、エピソード初期に C エージェントが T を獲る D エージェントに即座に侵襲され、C クラスタを形成することができず、拡大していくことが出来ない (図 5.4.2(a))。しかし混合戦略の場合は多くのエージェントの戦略が確率的に切り替わるため、D 型のエージェントのみが一方的に優位になることはない。初期のステップを耐え抜くことができればいくつか C 型クラスタを形成することができ最終的にはある程度の協調を維持できる (図 5.4.2(b))。つまり、混合戦略を用いることで、初期の D 型エージェントが拡大していき、C 型クラスタが D 型エージェントからの搾取に耐える必要がある Enduring Period (END)[108]において全員 D となる相に吸引される可能性が抑制されているといえる。また、混合戦略の場合、戦略適応、すなわち利得計算をシンクロに行うかアシンクロに行うかは協調の進化に大きな影響を与えない。利得計算をどのように行うかというのは、ゲーム中に確率的要素をどの程度加味するかということである。混合戦略 2×2 ゲームでは戦略決定を各エージェントの持つ戦略値に応じた確率で行うため、ゲーム中に大きな確率要素がすでに存在しており、利得計算をどのように行うかということによって生じるゲームプロセス中への確率要素の増減は最終的な結果に対し大きな影響を持たない。一方離散戦略定義の場合、利得計算をアシンクロに行うことで全体的に協調率はやや減少するものの、Dr が強くなった領域で協調がわずかにエンハンスされる。離散戦略 2×2 ゲームの場合、ゲームプロセス中に存在する確率要素は、今回の実験条件の場合、利得計算方法のみに影響される。離散戦略定義の場合、利得計算がアシンクロに行われることでゲームに確率的要素が加味され、混合戦略の結果同様、これによって D エージェントが一方的に拡大することを防ぎ、C エージェントの拡大がサポートされるのである。離散戦略でアシンクロに利得計算を行った場合や、混合戦略 2×2 ゲームのように確率的要素をゲームプロセスに加味すると、強ジレンマ領域である程度の協調が維持できるが、完璧な協調状態へと到達しにくくなる。これはゲームに介在するランダムネスが大きくなるため、完全に C エージェントが生き残るとは限らなくなるためである。このように、ゲームプロセスにおける確率要素の影響は協調の進化に大きな影響を及ぼすことが分かった。

次に、混合戦略定義での 2×2 ゲームでの結果の違いを詳しく見ていく。混合戦略定義は、自身の行動を戦略値に基づいて決定する際に、二通りの行動更新タイミングが想定できる。ここでは、混合戦略定義において行動更新ルールの違いは戦略進化の帰結にどう影響するかを見ていく。また、先の検討で、混合戦略 2×2 ゲームでは戦略適応すなわち利得計算がシン

クロかアシンクロは大きな影響がないことが示されたため、以降戦略はシンクロに更新することとする。

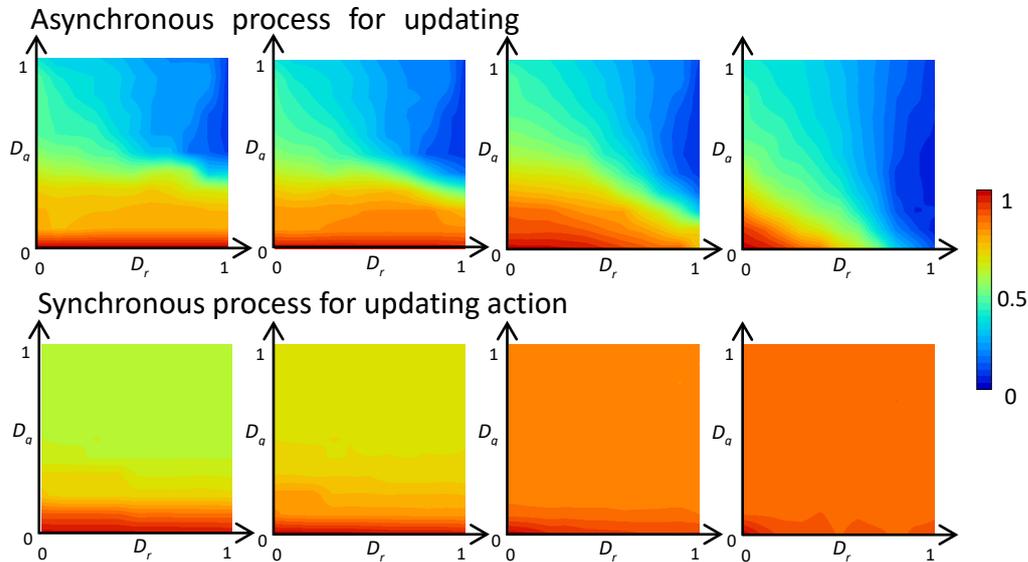


図 5.4.3 混合戦略モデルにおける均衡協調率. 次数は左から右に $k = 8, 12, 24, 48$ である. 上のパネルはアシンクロ行動更新, 下のパネルはシンクロ行動更新の時の均衡協調率である. 戦略更新はすべてシンクロ更新としている.

図 5.4.3 にアシンクロ的行動更新ルール, シンクロ的行動更新ルールそれぞれの時の結果を示す. 明らかに, 行動更新ルールの違いによって結果に大きな違いが表れていることが見て取れる. 特に, シンクロ的行動更新ルールを用いた場合, アシンクロ的行動更新ルールを用いた場合よりも協調が著しく **enhance** されている. ここで特筆すべき点は, アシンクロ的行動更新ルールを用いると, 平均次数が大きくなるほど協調が **enhance** されるという点である. ネットワーク互惠の本質は, ネットワークを仮定することで限定した繋がりを作り出し匿名性を減少させること, すなわち社会粘性の増加である. 匿名性の増加は **well-mixed** の状態への漸近を意味し, PDG 領域においては系全体を裏切り戦略へと吸引させることは明らかである. 従って, 次数の増加に伴って協調率が強ジレンマ領域において維持できなくなるとい現象が観察される **Asynchronous offer-update** の結果は離散戦略定義の 2×2 ゲームで予測される結果と類似の傾向を示しているといえるが, **Synchronous offer-update** の結果はそれとは相反するものであり, 極めて興味深いものであると言える.

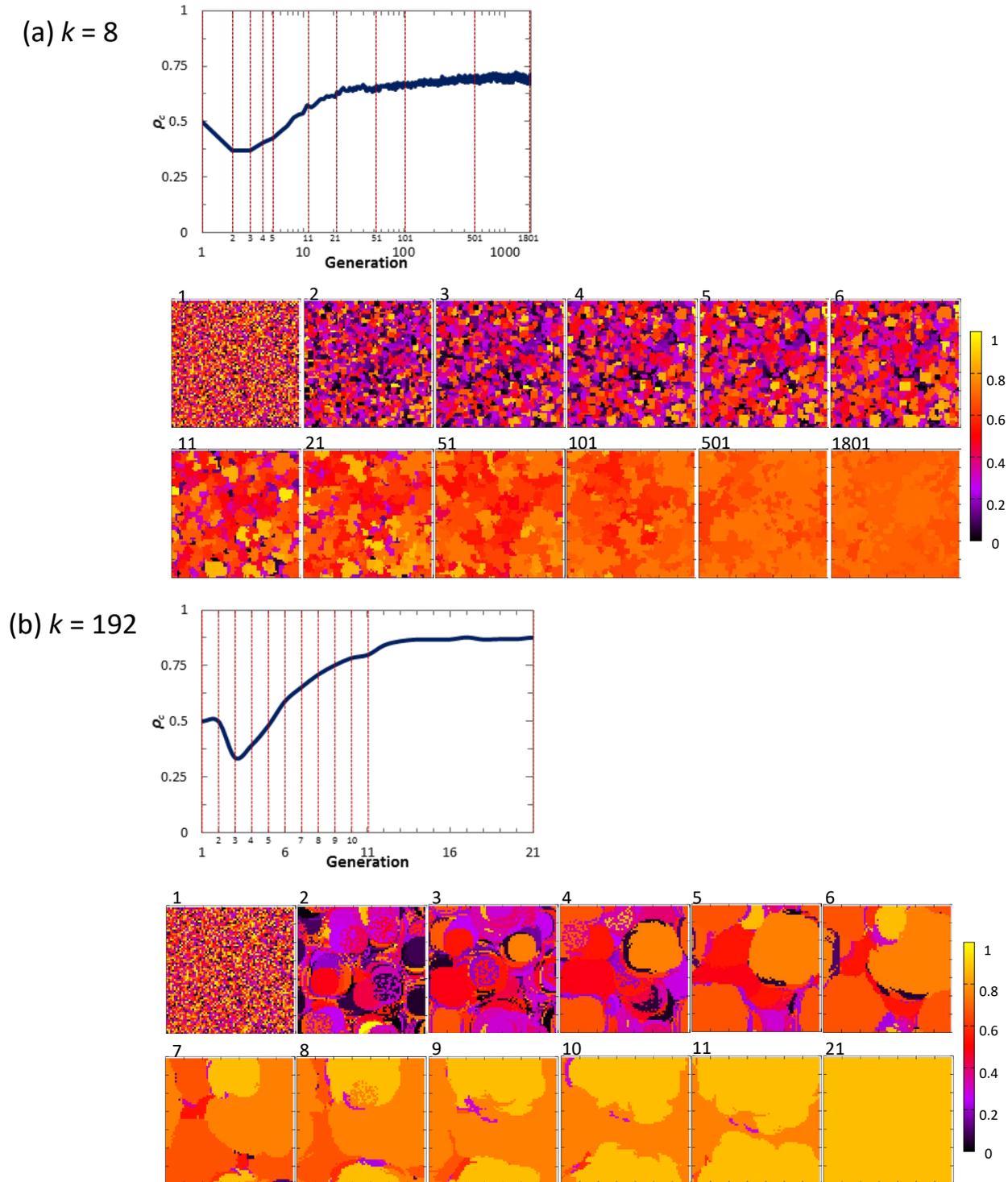


図 5.4.4 (a) $k = 8$ と (b) $k = 192$ のときの各エージェントの保有する戦略値 s_i のスナップショット。シンクロ行動更新を用いている。次数が大きいほど拡大速度と均衡協調率が高い。

混合戦略を用いる場合、どうして行動更新ルールがシンクロかアシンクロかという事が最終的な協調の進化に大きな影響を与えるのだろうか。スナップショットを見て詳しく考察し

ていく．図 5.4.4 に示す，regular lattice, $k=8$ の時のシンクロ的行動更新ルールの時のスナップショットを見ると，初期に D 型エージェントが C 型エージェントを侵襲して行く様子が分かる．そして，生き残った C 型クラスターが盛り返し，広がっていく．（このように，初期に著しく協調率が減少した後再び上昇に転じるという協調進化のメカニズムは Shigaki ら[108] が報告した END と EXP に関する報告と類似性がみられる（図 5.4.5）．初期に攻め込まれた C エージェントがクラスタリングに成功して広がっていくという現象は離散戦略定義のみならず，混合戦略定義 2×2 ゲームにおいても観察される普遍的な現象であることがわかる．）

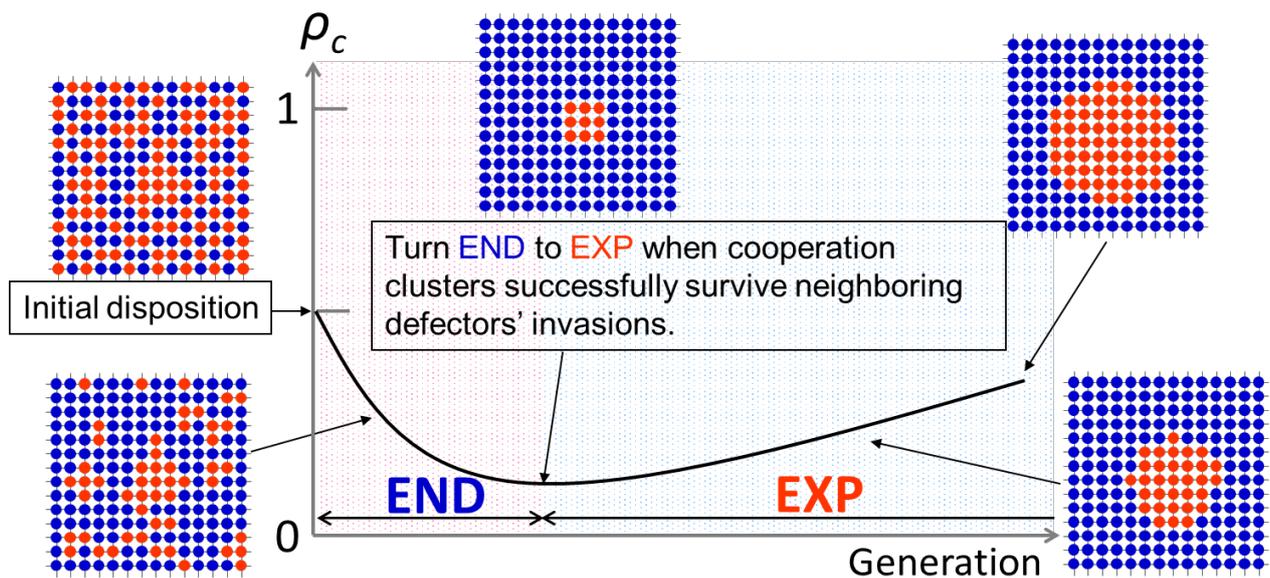


図 5.4.5 Shigaki ら[108]が報告した Homogeneous なネットワーク上における協調創発の普遍的枠組み．進化過程において初期に Enduring 期間(END)と呼ばれる協調率が大幅に減少する期間が存在する．それを耐えしのいた後，Expanding 期間(EXP)と呼ばれる Phase に突入し，協調クラスターが拡大する．

この時，図 5.4.6(a1)に示すエージェントの保有する戦略値と獲得利得の関係を見ると，エピソード初期においては，保有戦略値と利得に明らかな差がないものの，エピソード後期では，戦略値の大きなより協調型エージェントのほうが比較的多くの利得を獲得しているという傾向が読み取れる．PD 領域では戦略値が高いクラスター群のほうが境界を除いて常に高利得を獲得できるため（5.4.1.参照），協調率が比較的高いクラスターが一度クラスタリングに成功すれば，安定に存在できるのである．エピソード初期には，保有戦略値と各エージェントの獲得する平均利得に大きな差異は生じないが，最大利得は，エピソードを通して保有戦略値によ

って明らかな差異が生じる(図 5.4.6(a2)). 次数が大きくなると, こういった傾向はより顕著に表れる (図 5.4.6(b1)(b2)).

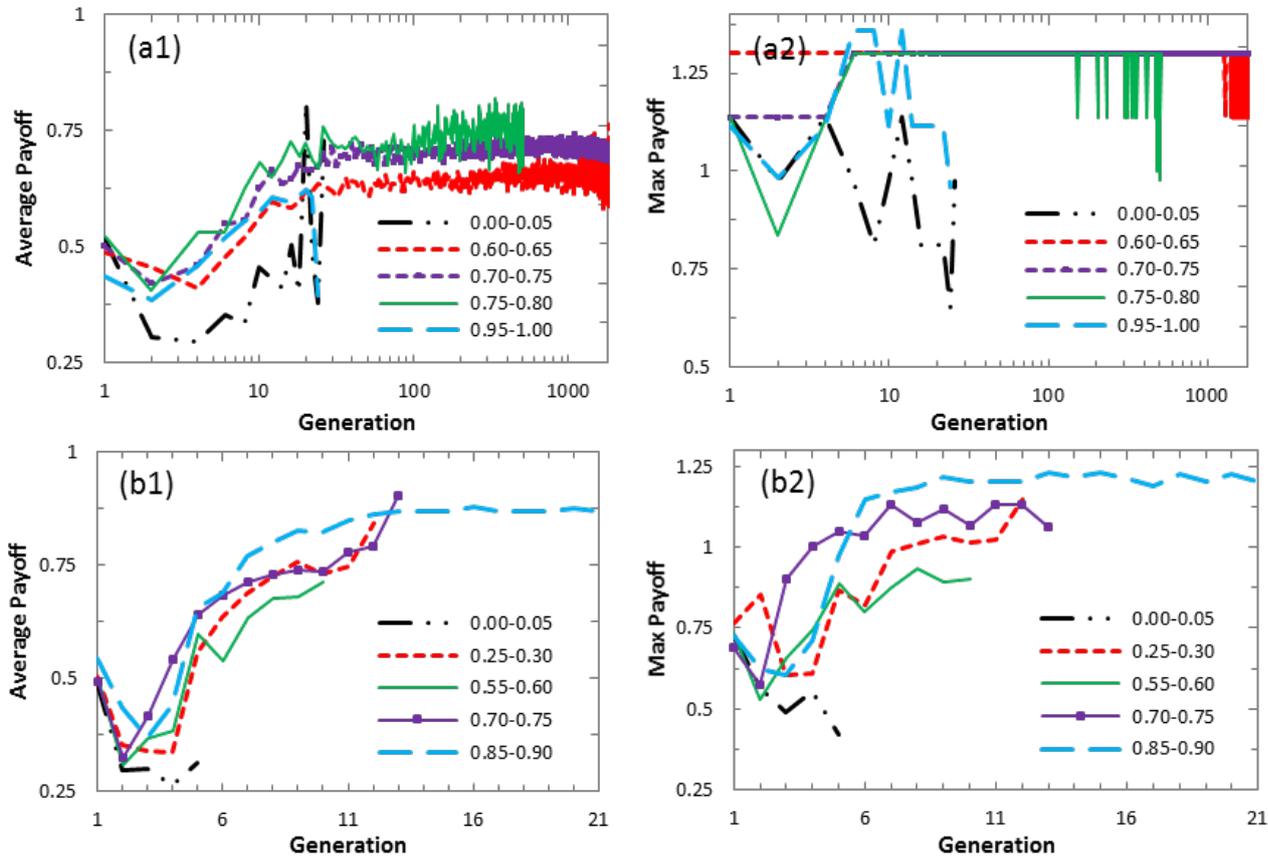


図 5.4.6 クラスターごとの時系列平均利得と最大利得. 図 5.4.4 に対応している. 上二つが $k=8$ のとき, 下二つが $k=192$ のときの図である. 次数が高くなると最大利得獲得者が明確化する.

次数が大きいとその分多くの隣人とゲームをする必要がある. 離散戦略定義の場合, 次数が大きい状態で PD ゲームを行うと, T を獲る裏切りエージェントが極めて高い利得を獲得し, 裏切り相へと容易に吸引される. 一方, 混合戦略ゲームでは, 各エージェントに付与された戦略値に応じた確率で戦略が各々決定される. すなわち, C 型のエージェントであるにも関わらず D 戦略を用いるという状況が不可避免的に生じる. この時, シンクロ的行動更新ルールを用いた場合, 協調型の戦略を保持しているにもかかわらず, ネットワークで規定された全隣人と D エージェントとしてゲームを行い, 比較的高い利得を獲得することが出来るエージェントが現れる. 平均次数が大きい場合, 戦略適応法として IM を用いていると, 次ステップで戦略が伝播する速度は非常に大きくなるため, 高利得を獲得した, D エージェントとし

て利得を獲得した協調型エージェントが安定なクラスターを即座に形成することができる。こういったメカニズムによって、シンクロ的行動更新の場合は高次数でも協調戦略が生き残ることができるのである。勿論平均次数が小さい場合も同様の現象は起こり得る。しかし、平均次数が小さいと、境界で接触するエージェントの影響を受けやすく、保有戦略値による明確な最大利得の差異が生じにくい。そのため、平均次数が大きい場合の時ほど協調がエンハンスされない。

混合戦略においてシンクロ行動更新を用いたとしても、際限なく協調が強化されるわけではなく、限界点が存在する。図 5.4.7 は基盤ネットワークの次数と均衡協調率の関係を表している。本枠組みを用いた場合、 $k=48$ 付近に協調率の最大値が出現する。これはジレンマ強さに依らず普遍的に観察されるピーク値である。しかしながら協調率が減少する速度は極めて緩やかであり、図 5.4.4 中で例示した $k=192$ のように大次数のケースであっても協調相に至る場合もある。ただ、そのような協調相が観察できる頻度は次数の増加に伴って徐々に減少する。

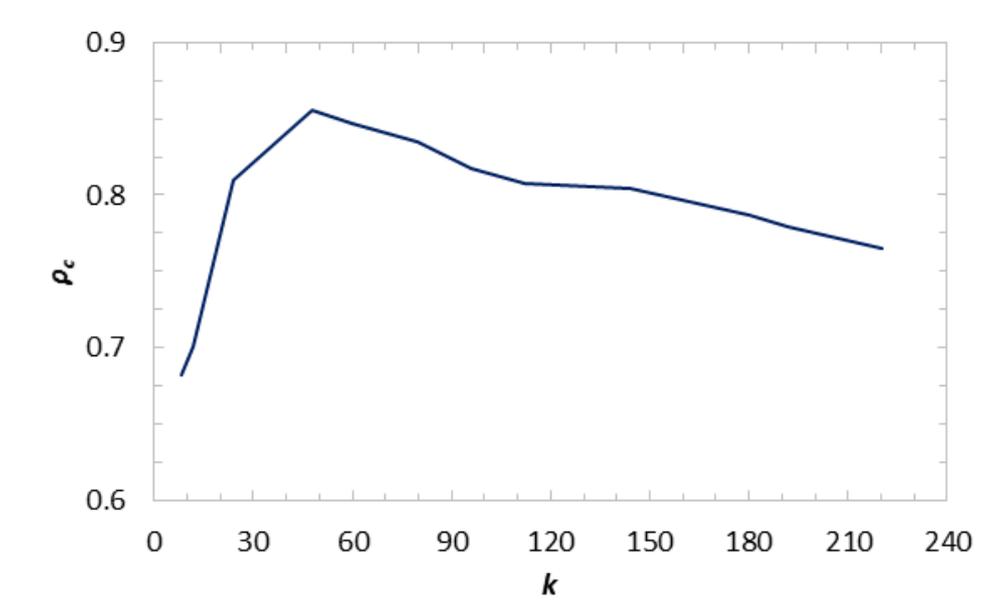


図 5.4.7 均衡協調率 ρ_c と基盤ネットワークの次数 k の関係。ジレンマ強さは $D_g = D_r = 0.3$ である。 $k=48$ 付近に均衡協調率のピーク値が出現する。

アシンクロ的行動更新の場合、フォーカルエージェントは隣人毎に手を再更新するため、シンクロ的行動更新を行う時のように劇的な利得の浮き沈みが起こりにくく、混合戦略であるということ以外特別に協調をエンハンスする効果はない。それ故シンクロ的行動更新より

も協調がエンハンスされない。行動更新をシンクロに行くかアシンクロに行くかということはゲームにおける確率的な要素をどの様に考慮するかの違いであり、系全体で見て、あるステップにおける C 戦略を採るエージェントと D 戦略を採るエージェントの比率がそれによって変わるわけではない。しかし本章での検討から、このような確率的要素の考慮の仕方の違いが最終的な結末に大きな違いを生じさせることが分かった。

5.4.1. (補足) 均衡協調率と平均利得の関係

クラスターサイズが十分に大きいとき、クラスターの中心にいるエージェントが得られる期待利得は平均場近似 (Mean Field Approximation) で導出が可能である。混合戦略 $s_x \in [0, 1]$ を保有するエージェント群のクラスターにおいて、予測される期待利得は

$$\begin{aligned} \langle \pi \rangle &= s_x^2 R + s_x(1 - s_x)s_x + (1 - s_x)s_x T + (1 - s_x)^2 P \\ &= (R - S - T + P)s_x^2 + (S + T - 2P)s_x + P \\ &= (R - S - T + P) \left[s_x + \frac{S + T - 2P}{2(R - S - T + P)} \right]^2 - \frac{(S + T - 2P)^2}{4(R - S - T + P)} + P \end{aligned} \quad (5.2)$$

s_x に関する微分をとると

$$\frac{d\langle \pi \rangle}{ds_x} = 2(R - S - T + P)s_x + S + T - P. \quad (5.3)$$

$\frac{d\langle \pi \rangle}{ds_x}$ が正であることを判定するために、 $R - S - T + P$ が正の値をとるか、0 か、負の値をとるかの場合分けする。 $\frac{d\langle \pi \rangle}{ds_x}$ の最小値が正であることを示すか、 $\frac{d\langle \pi \rangle}{ds_x} \Big|_{s_x=0}$ または $\frac{d\langle \pi \rangle}{ds_x} \Big|_{s_x=1}$ が正であることを

を示せばよい。以下に書く場合における正負を判定する。

(i) $R - S - T + P > 0$ のとき、

$$\begin{aligned} \frac{d\langle \pi \rangle}{ds_x} \Big|_{min} &= \frac{d\langle \pi \rangle}{ds_x} \Big|_{s_x=0} \\ &= 2(R - S - T + P)0 + S + T - P \\ &= S + T - P; \end{aligned}$$

(5.4)

(ii) $R - S - T + P = 0$ のとき、

$$\frac{d\langle \pi \rangle}{ds_x} \Big|_{min} = 2(0)s_x + S + T - P$$

$$= S + T - P; \quad (5.5)$$

(iii) $R - S - T + P < 0$ のとき,

$$\begin{aligned} \left. \frac{d\langle \pi \rangle}{ds_x} \right|_{min} &= 2(R - S - T + P)1 + S + T - P \\ &= 2R - S - T + P. \end{aligned} \quad (5.6)$$

以上より、 $\langle \pi \rangle$ が $0 \leq s_x \leq 1$ の範囲において単調増加関数であると言える。この条件はすなわち

$$2R > S + T - P > 0 \quad (5.7)$$

を満たしている。従って $R = 1$, $P = 0$ の一般的な PD ゲームを仮定し、 R 互恵が優位な領域のみを対象とする限り、上記の条件は常に満足されると言える。

また、 ST 互恵が優位になる場合、平均利得は s_x に関する単調増加関数とはならないため注意が必要である。

5.4.2. (補足) クラスタ中に生じる特殊な状態

図 5.4.8 に示すような、戦略 s_x を保有するクラスタの中心のエージェントのみが D 戦略を選択する特殊な状況が生起する確率を考える。このような特別な状況が生起する確率は平均場近似を用いて導出可能であり、

$$P_{\langle \pi \rangle max} = (1 - s_x)s_x^k \quad (5.8)$$

となる。ここで s_x はクラスタを構成するエージェントの保有する戦略、 k は次数 (クラスタ構成サイズ) を表す。これを図示したのが図 5.4.9 である。次数が大きくなるほど特殊な状況を生起し得る s_x が限定されてくる。つまり次数が $k=8$ のように小さいとあらゆる s_x で最大利得を供出し得る特殊な状況が生起するが、次数が大きくなると高い s_x を保有するクラスタがより特殊な状況を生起させやすい。これによって高い協調率に到達する試行が観察されるのである。

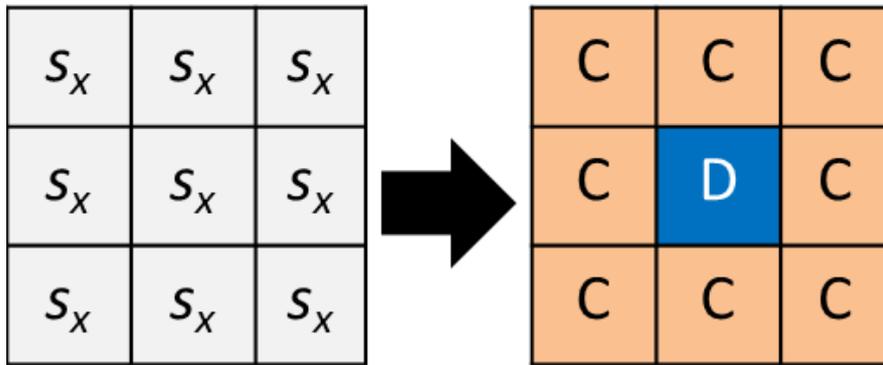


図 5.4.8 クラスターの中心エージェントのみ D 戦略をとる特殊な状況の模式図.

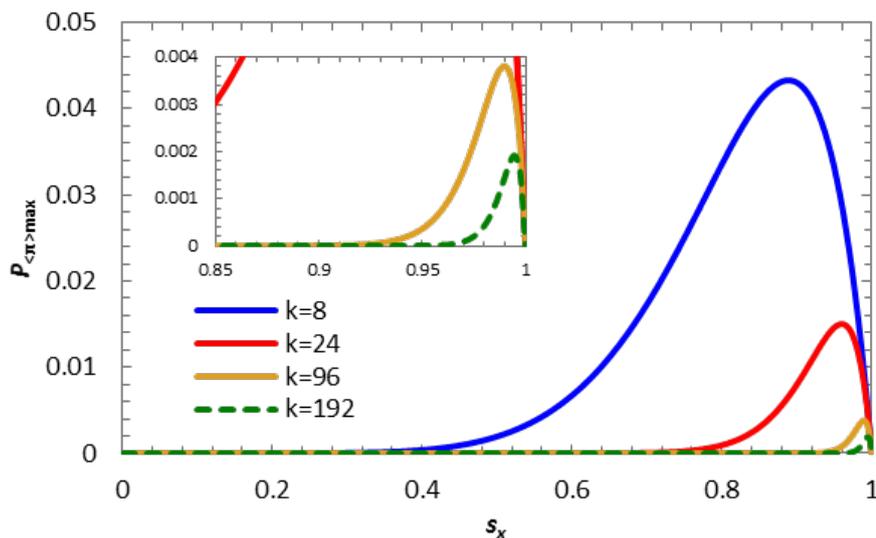


図 5.4.9 図 5.4.8 に図示するような特別な状況が生起する確率とクラスターの戦略 s_x の関係. クラスターサイズが $k=8, 24, 96, 192$ の 4 パターンを示す.

5.5. 結論

混合戦略定義 2×2 ゲームにおける協調創発メカニズムについて詳細な検討を行い、分散戦略定義の場合と比べて結果と協調創発機構にどのような違いが生じるかを考察した。特に、混合戦略 2×2 ゲームにおいては各エージェントがどのタイミングで戦略値に基づいた行動の更新を行うかの行動更新ルール規定が、最終的な結果に劇的な違いをもたらすことを示した。とりわけ、シンクロ的行動更新ルールを用いた場合、協調は劇的にエンハンスされ、かつ平均次数が大きくなるほど協調がエンハンスされるという結果となった。シンクロ更新とアシンクロ更新の違いは、ゲームプロセス中への確率要素の考慮の仕方の違いである。分散戦略ゲームでは、戦略更新がシンクロかアシンクロかの差異で生じる確率性要素の考慮の仕方

による影響は、比較的大きいことが知られている。しかし、本章で示しているように、混合戦略ゲームでは、戦略それ自体に確率性が強く加味されており、戦略更新がシンクロかアシンクロかによる確率的要素の混入具合は結果に大きな影響を与えない。こういった事実があるにも関わらず、混合戦略ゲームにおいて、どのタイミングで自身の行動を決定するかを規定することで生じる確率的な揺らぎは、最終的な協調の進化に大きく影響を与えることが分かった。このことから言えるのは、現実社会において極わずかと思える人の行動の揺らぎでさえも、複雑に絡み合った社会に対して、非常に大きな影響を与えうるという事である。

6. 中間（日和見）戦略の定義による囚人のジレンマ解消メカニズムの解明

6.1. 研究背景と着眼点

混合戦略ゲームは前章で詳述したとおり，多様な人の行動様式をゲームモデルに再現しうる画期的な方法の一つである．混合戦略ゲームにおいて，各エージェントは $s_i \in [0, 1]$ の実数値を付与され， s_i に基づいた確率で次手 C（協調）または D（裏切り）を決定する．混合戦略による確率的な要素が協調の創発に寄与していることがわかっているが，本章では混合戦略として付与される実数 s_i をより詳細に分類し離散戦略と完全な混合戦略の中間的なふるまいをする戦略を定義することにより，混合戦略における詳細な協調創発メカニズムをより深く分析する．

6.2. 研究概要

混合戦略ゲームでは各エージェントは保有戦略値 s_i に応じた確率で次回の行動を決定する．本章では付与される実数値の戦略解像度 N_R と協調の進化の関係を検証する．具体的に $N_R=2$ の場合離散戦略定義を表し， $N_R=\infty$ の場合混合戦略定義を意味する．本章では，囚人のジレンマゲームとチキンゲームにおいて， N_R の違いによって協調の進化の様子が大きく異なることを発見した．また， $N_R=3$ のとき，特に強い協調エンハンス効果を確認した．

6.3. モデル

6.3.1. 2×2 ゲーム

2×2 ゲームを基盤モデルとして用いる．混合戦略定義ではエージェント i は 0 から 1 の間の実数値で定義された戦略 s_i を保有する．本章では s_i はエージェント i が協調行動（C）をとる確率を表す．従って， $s_i=1$ のとき，エージェント i は常に協調する協調者，一方 $s_i=0$ のとき常に裏切る裏切者を表す． $D_g = T - R$ ， $D_r = P - S$ とし，簡単のため $R=1$ かつ $P=0$ とすることで利得表は

$$\begin{pmatrix} R & S \\ T & P \end{pmatrix} = \begin{pmatrix} 1 & -D_r \\ 1 + D_g & 0 \end{pmatrix} \quad (6.1)$$

と書き表せる．本章では代表的なジレンマゲームである， $T > R > P > S$ かつ $2R > T + S$ を満

たす $0 \leq D_g \leq 1$ かつ $0 \leq D_r \leq 1$ である囚人のジレンマゲームと、部分的に $2R < S + T$ を満たす $0 \leq D_g \leq 1$ かつ $-1 \leq D_r \leq 0$ であるチキンゲームを仮定する。

本章では混合戦略値の多様性を表す指標として戦略解像度 N_R を定義する。 $N_R = 2$ のとき $s_i = 0$ または 1 のみが許容される完全な離散戦略を、 $N_R = \infty$ のとき、 $s_i \in [0, 1]$ である完全な混合戦略ゲームを表す。但し、本章の使用 Platform の関係から $N_R = \infty$ は $N_R = 32768$ と同義とする。ここで s_i と N_R の関係を整理すると、

$$s_i \in \left\{ \frac{0}{N_R-1}, \frac{1}{N_R-1}, \frac{2}{N_R-1}, \dots, \frac{N_R-1}{N_R-1} \right\} N_R \geq 2 \quad (6.2)$$

となる。

6.3.2. シミュレーション手順

最初に次数 8 の二次元格子状に各エージェントは配置される。本章ではそうエージェント数 $N = 70 \times 70 = 4900$ とする。各エージェントはネットワークで繋がった隣人とゲームを行い利得を獲得する。混合戦略ゲームにおいて自身の次手 C または D を決定するための行動更新ルールは 2 種類仮定できる。本章では前章で説明した、アシンクロ行動更新を行動更新ルールとして用いる。すなわち各エージェントは隣人ごとに戦略値 s_i に応じた確率で協調するか裏切るかを毎度決定する。利得が積算されたのち、戦略更新を行う。略更新ルールとしては、隣人中の最大利得者の戦略をコピーしてくる **Imitation Max(IM)** を用いる。戦略はシンクロに更新される。このプロセスを均衡に至るまで繰り返し、均衡時の協調者の人数である均衡協調率を考察する。統計的有意性を確保するため、ジレンマ強度ごとに 20 試行平均した協調率を結果として用いる。

6.4. 結果と考察

図 6.4.1 は囚人のジレンマゲームにおける均衡協調率を表す。戦略解像度 N_R が大きくなるにしたがって D_g , D_r の大きな強ジレンマ域においてもある程度の協調が維持できるようになる。しかしながら、 N_R が大きくなると協調率 ρ_c は完全に 1 に至るわけではなく、中途半端な値で均衡する。特にチキンゲームの性質が色濃く残る D_g が大きく、 D_r の小さい領域で、その傾向は顕著である。チキンゲームにおける Nash 均衡解は C と D 併存平衡である。従って、混合戦略が許容されることでチキンゲームにおける安定点である併存状態（中途半端な均衡協調率）に至ったと考えられる。

$N_R = 3$ のとき、stag-hunt (SH) ゲームに近い部分で大きな協調エンハンス効果が確認でき

る。また、それ以外の場合においても SH ゲームとの境界線上においてはほぼ完全な協調相を達成している。このことから、戦略の多様性が増加すると、 D_r に対して協調率は感度が小さくなると言える。

N_R の違いがなぜ均衡協調率に影響を与えるのかを知るため、スナップショットを見る。図 6.4.2 は $D_g = 0.1$, $D_r = 0.7$ におけるスナップショットを表す。 $N_R = 2$ の場合すなわち離散戦略定義の場合を見ると、即座に裏切りエージェントが系を支配し、協調者は数個の極めて小さなクラスターが残るのみとなる。

$N_R = 3$ の場合を見てみると、 $s_i = 0.5$ のエージェントが $s_i = 0$ の裏切りエージェントと $s_i = 1$ の協調エージェントと共にクラスターを形成する。このとき C クラスターは $N_R = 2$ の離散戦略定義のときとほぼ同様の形態で生存する。しかしながらここから $s_i = 0.5$ を保有するランダムに振る舞うエージェントが C クラスターと D クラスターの間に入り込むことで、C クラスターが拡大に転じる。一度 C クラスターが拡大に転じると D クラスターを侵襲し、協調相を達成することができる。 $s_i = 0.5$ を保有するランダムに振る舞うエージェントは時に協調戦略を選択する場合がある。これによって境界部の C エージェントの利得が増加し、C クラスターの拡大に寄与するのである。図 6.4.3 を見ると $s_i = 0.5$ の日和見エージェントは初期に D 型エージェントと比べて比較的大きな利得を獲得し、D クラスターの拡大を阻害していることが見て取れる。

$N_R = 4$ のときを見てみると、図 6.4.1 に示している通り広いジレンマ領域である程度の協調エンハンス効果があるものの、完全な協調相には至り得ない。図 6.4.2 のスナップショットを見てみる。 $N_R = 4$ のとき、戦略値のバリエーションは $s_i \in \{0, 0.3333, 0.6666, 1\}$ である。初期には D の侵襲を受け、戦略値ごとにまとまって形成されたクラスターが生成される。いったん $s_i = 1$ を保有する完全な C クラスターが拡大に転じる。しかしながら $s_i = 0.6666$ を保有するクラスター群に侵襲されてしまう。これは $s_i = 0.6666$ を保有するクラスターの方が T を獲得する分獲得利得が完全 C クラスターよりも大きくなるためである。

$N_R = 5$ の場合を見てみると他の 2 ケースと同様に戦略値ごとにクラスターを形成して D の侵襲を耐える。その後 $s_i = 1$ を保有する完全な C クラスターは拡大に転じるものの、前述の理由で最終的には $s_i = 0.75$ の協調型エージェントが系を支配する。

戦略解像度 N_R が大きくなると完全な C クラスターが崩壊し、生存できなくなるものの、比較的協調的なクラスターは生き残れるようになる。 N_R が 30 以上になるともはや解像度が均衡に与える影響は極めて小さくなる。

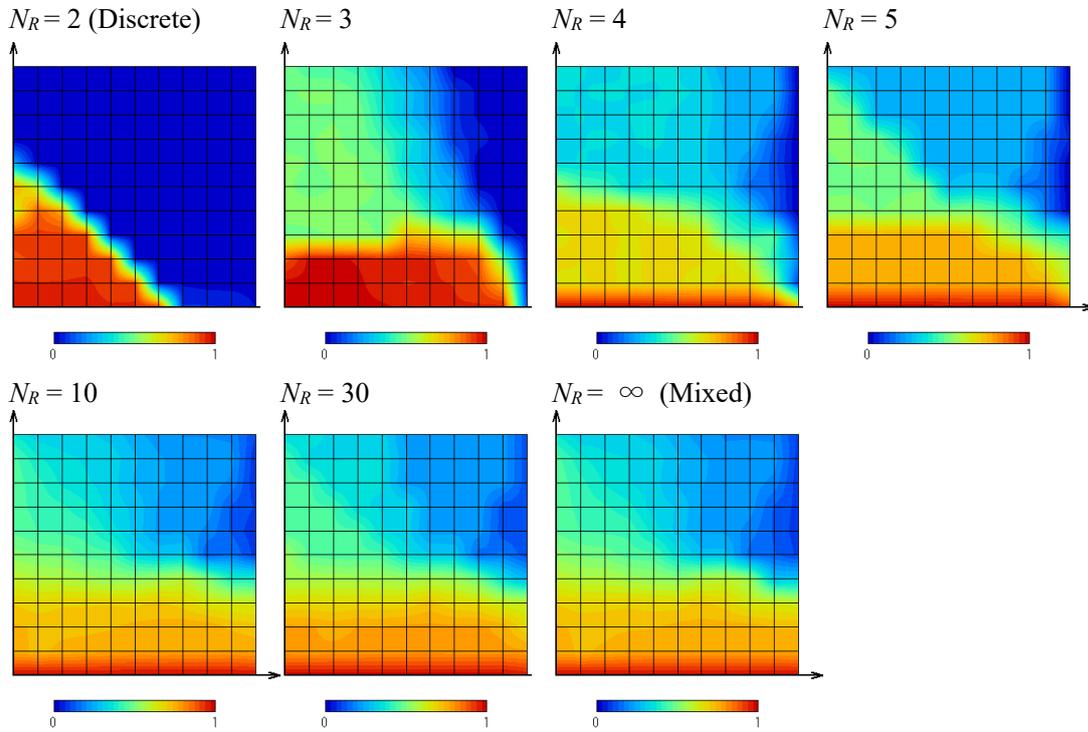


図 6.4.1 囚人のジレンマゲームにおける均衡協調率 ρ_c . N_R は戦略解像度を表しており, $N_R=2$ のとき完全な離散戦略定義を, $N_R = \infty$ のとき完全な混合戦略定義であることを表す. N_R が大きくなるほど強ジレンマ域でもある程度の協調率が維持できるようになる. しかしながら, 完全な協調相には達し得にくくなる. $N_R=3$ のとき, stag-hunt ゲームとの境界付近で大きな協調エンハンス効果が観察される.

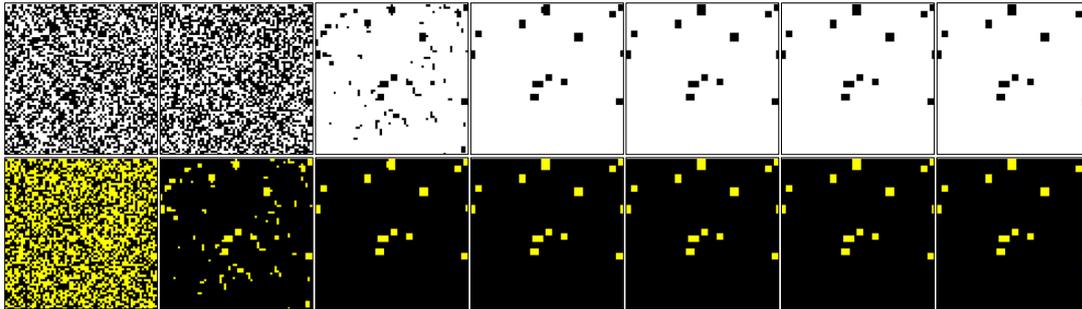
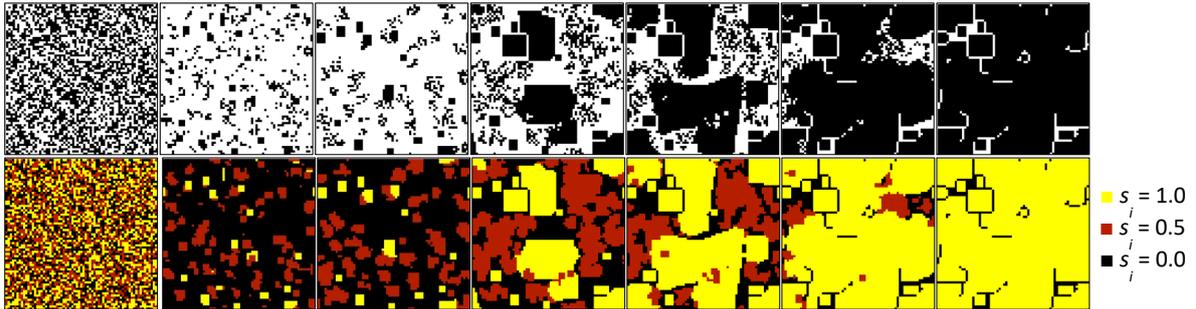
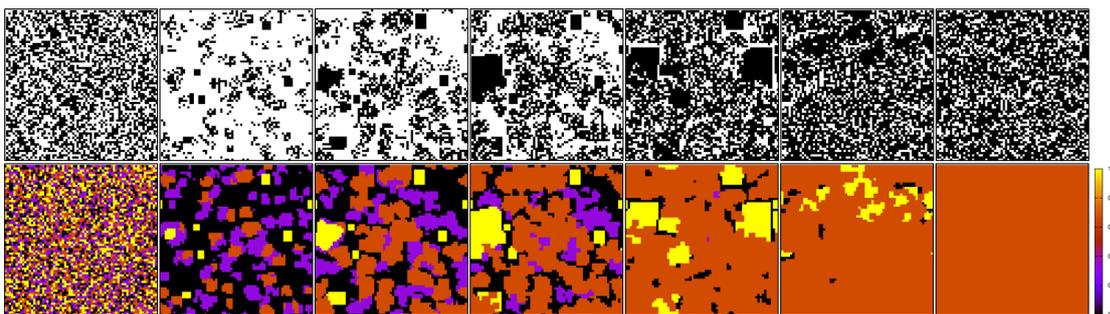
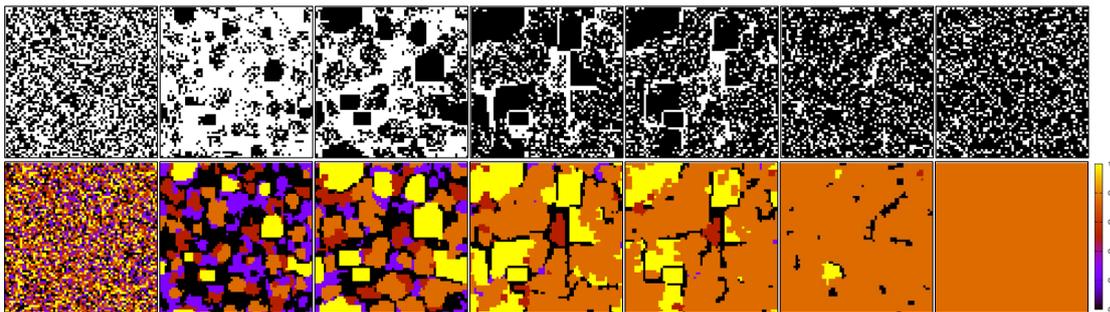
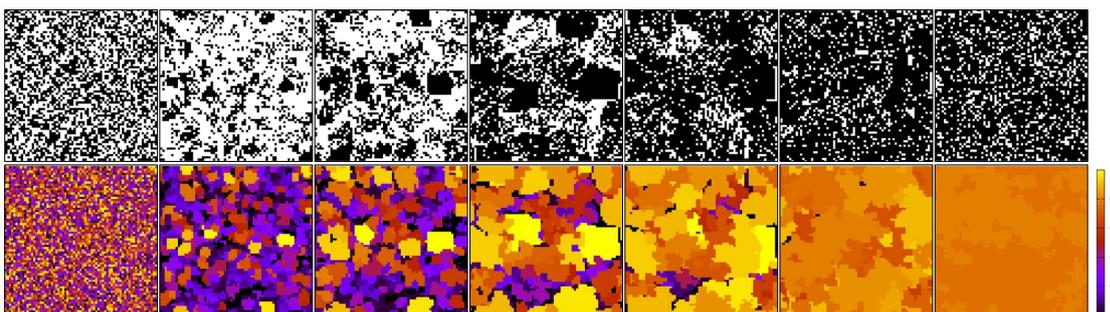
$N_R = 2$

 $N_R = 3$

 $N_R = 4$

 $N_R = 5$

 $N_R = 30$


図 6.4.2 $D_g=0.1$, $D_r=0.7$ におけるスナップショット. 離散戦略定義である $N_R=2$ のとき C クラスタは D に侵襲され少数が極めて小さなクラスターを形成し生き残るに過ぎない. しかしながら $N_R=3$ のとき $s_i=0.5$ を保有するランダムに振る舞うエージェントが境界部の C エージェントの利得増加に寄与し, C クラスタが拡大する. N_R が大きくなると完全な C クラスタが協調型クラスターに侵襲されるため, 完全な協調相には至らない.

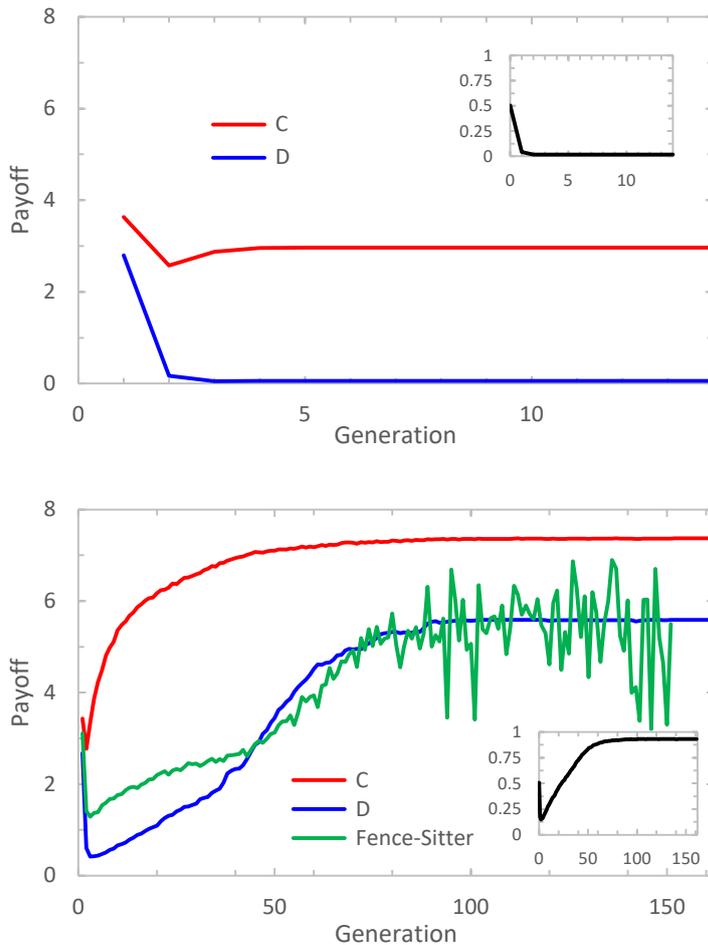


図 6.4.3 $D_g=0.1$, $D_r=0.7$ における時系列平均利得 (図 6.4.2 と対応). 差し込み図は平均協調率を表している. 上図は $N_R=2$, 下図は $N_R=3$ である. C とは完全協調型エージェント ($s_i=1.0$), D とは完全裏切エージェント ($s_i=0$), Fence-Sitter とは日和見型エージェント ($s_i=0.5$) を表す.

図 6.4.2 をみるとチキンゲームとの境界付近でかつ D_g が大きくなると $N_R = 2$ よりも $N_R = 3$ のほうが協調率が減少する．この理由を考察するため， $N_R = 2$ よりも $N_R = 3$ のほうが協調率が上昇する領域との比較を行う．図 6.4.4 は $D_g = 0.3$ ， $D_r = 0.1$ におけるスナップショットである．これを見ると， $s_i = 0.5$ のエージェントが存在する場合，協調クラスターの拡大の後最終的に $s_i = 0.5$ の日和見型エージェントに系が支配される様子が分かる．図 6.4.5 の時系列データの比較から， N_R によらずエピソード序盤は C の平均利得は常に D や Fence-Sitter ($s_i = 0.5$ を保有) のものを上回っている．しかしながら $N_R = 3$ のとき最終的に Fence-Sitter に系が支配される．この理由をさらに詳しく検証するために，図 6.4.6 の最大利得時系列を見る．これを見ると $D_g = 0.1$ ， $D_r = 0.7$ のとき，Fence-Sitter の最大利得が大きく C や D のものを上回っている．これによって Fence-Sitter は D を駆逐すると同時に C も駆逐してしまうのである．

図 6.4.7 の最低利得を示す．特に際立った変化はないが， $D_g = 0.3$ ， $D_r = 0.1$ のケースの場合，戦略群間の最低利得差が $D_g = 0.1$ ， $D_r = 0.7$ と比べて若干小さい．

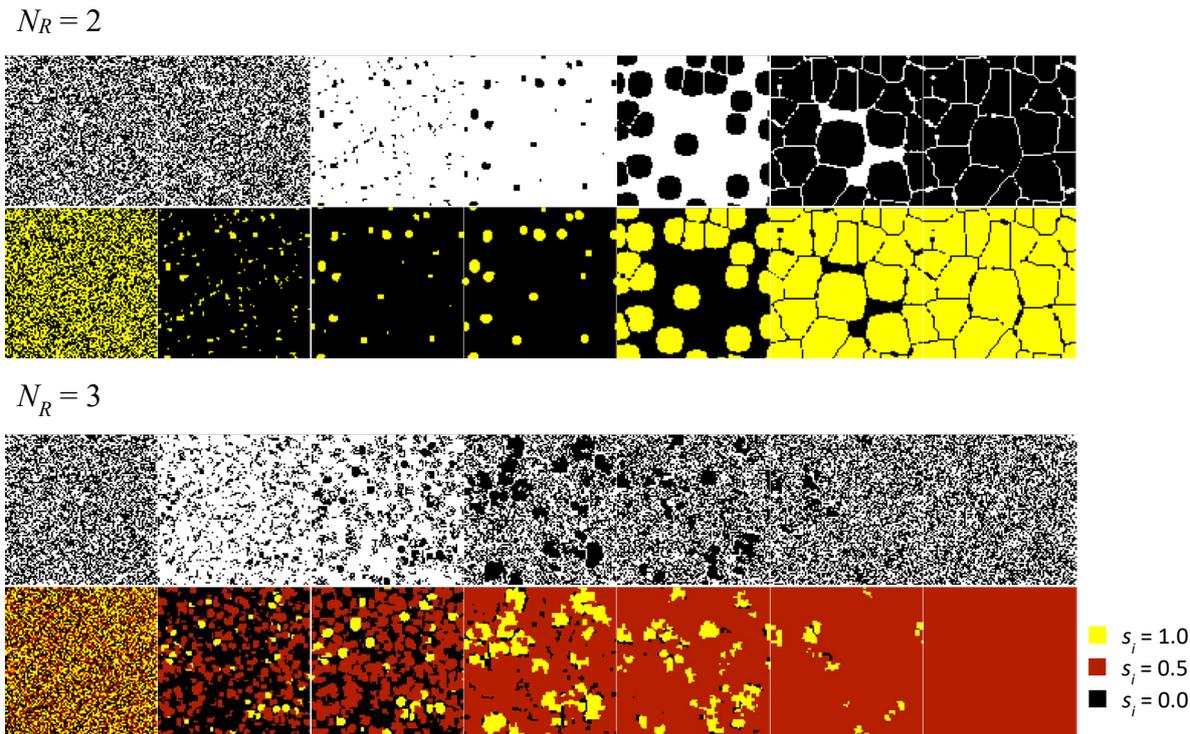


図 6.4.4 $D_g = 0.3$ ， $D_r = 0.1$ におけるスナップショット．分散戦略定義である $N_R = 2$ のとき初期の D の侵襲に耐えた細かい C クラスタが拡大していく．一方 $N_R = 3$ のとき，生存した C クラスタは一時拡大するものの，最終的には $s_i = 0.5$ を保有する日和見エージェントに系を支配される．

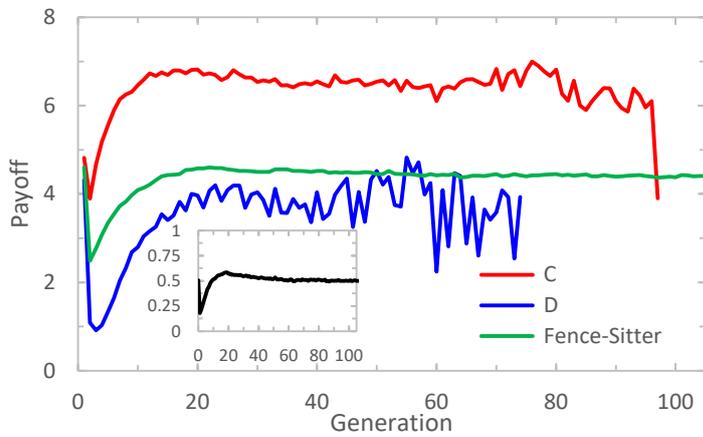
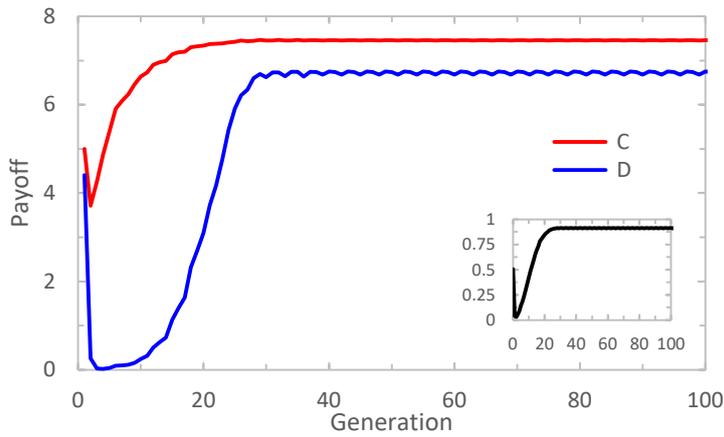


図 6.4.5 $D_g=0.3$, $D_r=0.1$ における時系列平均利得 (図 3 と対応). 差し込み図は平均協調率を表している. 上図は $N_R=2$, 下図は $N_R=3$ である.

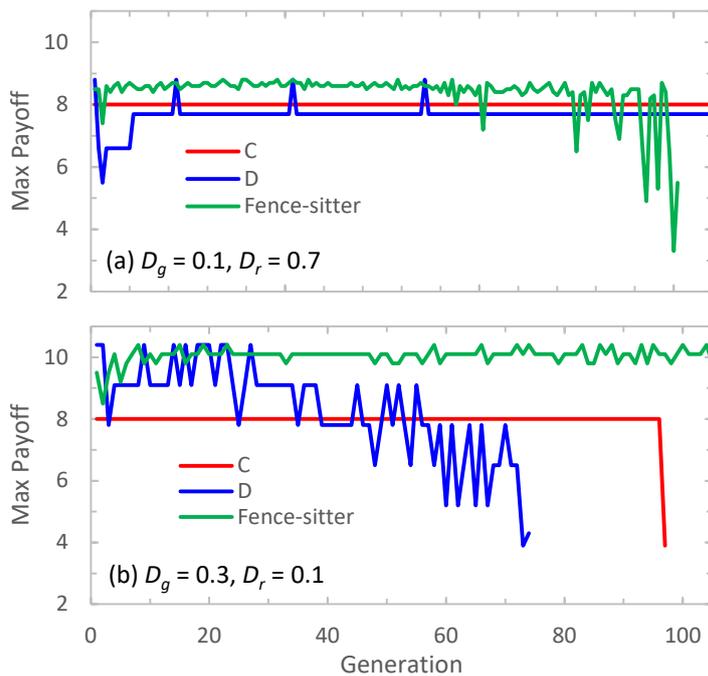


図 6.4.6 時系列最大利得. (a) $D_g = 0.1$, $D_r = 0.7$, (b) $D_g = 0.3$, $D_r = 0.1$ で $N_R = 3$ のときある. 凡例の Fence-Sitter (日和見エージェント) とは $s_i = 0.5$ を保有数エージェントの最大利得のことである.

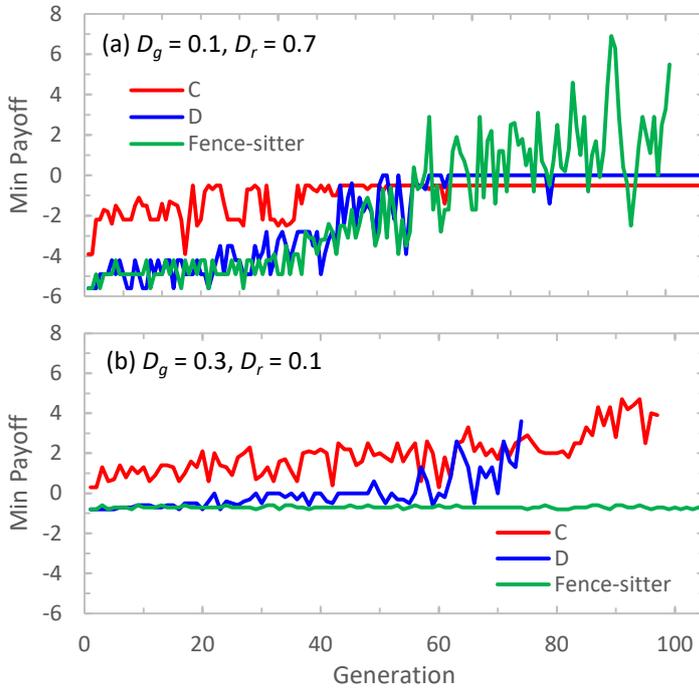


図 6.4.7 時系列最低利得. (a) $D_g = 0.1$, $D_r = 0.7$, (b) $D_g = 0.3$, $D_r = 0.1$ で $N_R = 3$ のときある. 凡例の Fence-Sitter (日和見エージェント) とは $s_i = 0.5$ を保有数エージェントの最小利得のことである.

6.5. 結論

混合戦略定義モデルにおいて戦略を付与する際、エージェントに与える実数戦略値の分割数を変えることで離散戦略から完全な混合戦略との中間の戦略地の定義し、協調の進化にどのような影響を与えるのかを検証した。この実験の結果、戦略解像度 N_R を 30 以上に設定すれば十分に混合戦略ゲームの特性が現れることが分かった。また、解像度を $N_R = 3$ に設定した時、Stag-hunt ゲームとの境界付近ほど強固な協調エンハンス効果が確認できた。このことから、協調の進化において $s_i = 0.5$ のようにどっちつかずの振る舞いをするエージェントが系に一定数存在することで協調クラスターの拡大がサポートされることが分かった。

混合的な特性をエージェントに付与すると強ジレンマ領域でもある程度の協調率が維持されるものの、最終的に至り得る均衡点は中途半端なものとなる。このような状態はある意味で現実的な結果と言えるのかもしれない。私たちの社会では完全な協調者または裏切者のように、常に一定の振る舞いをしていると言えれば言い過ぎである。この見地から、隣人によって振る舞いを多様に変える可能性がある混合戦略によるシミュレーションもまた一つの現実的

状況である。この実験の結果から、振る舞いの多様性が協調の進化に関連していることが分かったが、今後はヘテロネットワークなどでより複雑な相互作用が生じる場面での協調の進化とジレンマの関係を検証する必要があるだろう。

7. ネットワーク上の同類的な繋がり の促進が囚人のジレンマ ゲームにおける協調進化に与える影響

7.1. 研究背景と着眼点

戦略のみならずネットワークを進化させる機構を共進化モデルと一般的に呼称する．ネットワークの進化を規定するにあたって，Zimmermann ら[6]や Tanimoto[109][110]，Santos ら[111]に代表されるように，自分と隣人の戦略の組み合わせや，利得の大小比較によってリンクの切断を決定するのが一般的である．本章では，これらの既往研究とは視点を変え，戦略の同類性に注目する．具体的には，戦略が同類である者同士のリンクを維持し，そうでないものとのリンクを切断するものである．本章では，この提案ネットワーク進化ルールによって同類的な繋がりを促進することが，囚人のジレンマゲームにおける協調の進化にどのような影響を及ぼすのかを検討する．

7.2. 研究概要

本章では，ネットワークと混合戦略の共進化モデルを提案する．混合戦略の場合，戦略 s_i は 0 から 1 の間の実数値で表記され，各エージェントは保有する戦略値 s_i に応じた確率で自分の振る舞いを決定する．そのため，混合戦略の場合エージェントが取り得る戦略は多様である．そこで本章では，このような多様なエージェントが存在する集団において，戦略の同類性と協調の進化の関連性に注目する．各エージェントは自分の保有する混合戦略と似通った戦略を持つエージェントとのみネットワークの接続を許容することで，戦略同類性の高いエージェント同士との繋がりを促進する．この実験の結果，同類的な繋がりを促進することで，より強固な協調が創発することを報告する．また，ネットワーク適応速度が相対的に遅い場合において，戦略適応の際に適切な確率でコピーエラーを生起させることがネットワークの不均質性を増加させ，協調がより進化することを報告する．

7.3. モデル

7.3.1. 2×2 ゲーム

2×2 ゲームモデルを用いる．提案モデルでは，実数値からなる混合戦略 $s_i \in [0, 1]$ を導入する．各エージェントは確率 s_i で協調戦略 (C) をとる．2人2戦略ゲームでは $s_i \in \{0, 1\}$ のよ

うにバイバリーに戦略を付与する離散戦略が一般的な戦略定義であるが、提案モデルでは人のランダムな振る舞いをモデル化するため、確率的な意思決定方法として実数値戦略を用いる。各エージェントは、戦略によって決定される自分と相手の振る舞いのに応じて利得を獲得する。もし、自分と対戦相手共に C 戦略を選択した場合、Reward (R) を、両者共に D を選択した場合、Punishment (P) をそれぞれ利得として得る。もし、一方が C、他方が D 戦略を用いた場合、前者は Sucker (S)、後者は Temptation (T) を得る。これら利得 R, S, T, P の大小関係を変化させることで、様々なジレンマ構造をモデル中に再現できる。ここで、先行研究 [103] に倣い $R=1, P=0$ と固定することで、

$$\begin{pmatrix} R & S \\ T & P \end{pmatrix} = \begin{pmatrix} 1 & -r \\ r & 0 \end{pmatrix} \quad (7.1)$$

とする。この操作によって囚人のジレンマゲームを $0 \leq r \leq 1$ の範囲で表現できる。

7.3.2. 共進化モデルにおけるシミュレーション

シミュレーションは以下に示す手順で行われる。各エージェントは初期に次数 $k=4$ の 2 次元正方格子ネットワーク (ラティスネットワーク) 上に隙間なく配置される。集団からランダムに選ばれたエージェント i はネットワークで繋がった隣人とゲームし利得を獲得する。その後、確率 w でネットワーク適応を行い、そうであれば戦略適応を行う。ここで w はネットワーク適応速度パラメーターであり、 w が大きいほどネットワークの進化速度が大きく、 $w=0$ の時、戦略のみが進化することを意味する。この適応プロセスを均衡に至るまで繰り返し、最終 100 ステップの平均協調率 ρ_c (集団中での C 戦略をとるエージェントの割合)、平均戦略値を獲得し、考察する。

7.3.3. ゲーム

ネットワーク上からランダムに選ばれたエージェント i はネットワークで規定された全隣人 j とゲームを行う。エージェント i は隣人 j と対戦するとき、確率 s_i で協調戦略 C をとり、対戦相手 j は確率 s_j で協調戦略 C をとる。確率で決定された各々の振る舞い C または D に応じて利得表に基づいた利得が積算される。全対戦でエージェント i は自身の持つ戦略に応じた確率で振る舞いを更新していくため、同一ステップ中の対戦であっても隣人 j_1 に対する振る舞いと隣人 j_2 に対する振る舞いが必ずしも一致するとは限らない。隣人 j も同様に各々その隣人 k とゲームを行い、利得を獲得する。すなわち本章では前章で混合戦略定義モデルの

ための行動更新ルールとして定義したアシンクロ行動更新とアシンクロ戦略更新ルールを用いる。

7.3.4 戦略適応

戦略適応が行われる場合、エージェント i とその隣人集団 j の利得が確定後、エージェント i は、隣人集団の中から最も高い利得を獲得している隣人 j を一人選択し、その戦略を無条件にコピーする (Imitation Max; IM)。

戦略をコピーするとき、確率 μ でコピーしてくる戦略が 0 から 1 の間の実数値にランダムに変異する。これをコピーエラーと呼ぶ。

7.3.5. ネットワーク適応—同類的接続

ネットワーク適応が行われる場合、エージェント i は隣人エージェント j を一人ランダムに選択する。もし $|s_i - s_j| > \alpha$ の場合、エージェント i は j とのリンクを切断し、現有隣人以外の新たな隣人と再接続をし、 j はリンクを 1 本失う。一方 $|s_i - s_j| \leq \alpha$ のとき、 i と j のリンクは保持される。ここで、 $\alpha \in (0, 1)$ は許容度パラメーター (tolerance parameter) である。 α が小さいほどエージェントは同類な隣人を好むようになり、非同類的な戦略を持つエージェント同士のリンクは排除されることで、結果的に同類的な繋がりを持つグループの形成が促進される。本章ではこの同類的な集団形成を促す接続機構を *Assortative attachment* と呼称する。

7.4. 結果と考察

本モデルにおける、戦略と assortative attachment によるネットワークの共進化が協調の進化にもたらす影響について詳しく考察していく。図 7.4.1 にネットワーク適応速度 $w=0.1, 0.5, 0.9$ それぞれの均衡協調率をそれぞれ示している。ここで、 r はジレンマ強さである。この結果を俯瞰すると、戦略とネットワークの共進化、すなわち assortative attachment を許可することによって従来の固定ネットワーク時よりも協調がエンハンスされていることが分かる。tolerance parameter ; α の影響に注目すると、 w が小さい時、協調率が急激に減少する臨界点 $r = 0.50$ 付近では α の差異が観察されるものの、取り立てて言及するほどの明らかな影響は見られない。しかし、 w が大きくなると、協調の進化に対して α に依る明らかな影響が観察される。その傾向は w が大きいとき、 α が小さいほど強いジレンマ域でも協調を維持できるというものである。パラメーター α による顕著な影響が現れるのは、 w が大きいときネットワー

ク適応の影響力が卓越するからである。これは、ネットワーク適応が早い頻度で行われるならば、自戦略との許容度を小さくして同類隣人が集まってクラスターを形成した方が、大きなジレンマに耐えられることを意味する。

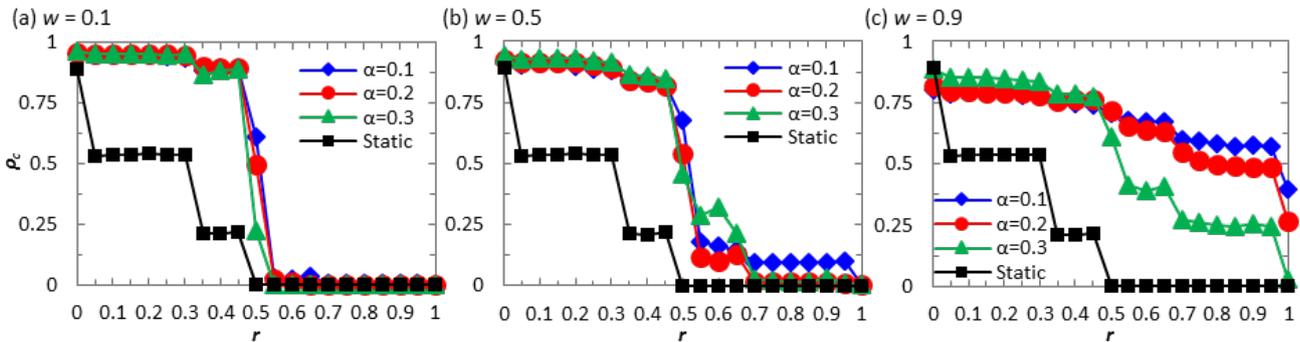
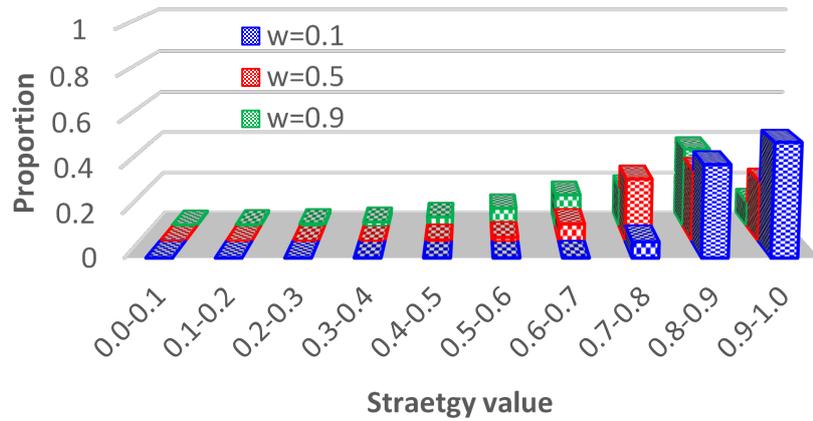


図 7.4.1 (a) $w = 0.1$, (b) $w = 0.5$ and (b) $w = 0.9$ それぞれの場合における均衡協調率 ρ_c . ”Static” は $w=0$ すなわち戦略適応のみ生起する状況の結果を表している。戦略適応時のコピーエラーは考慮されていない。

本モデルでの assortative attachment がなぜ協調を促進するのか、さらに詳しく考察していく。そもそもネットワーク進化という機構は、ある条件に従って自分にとって不利な interaction を排除・再構成するいわゆる Game exit option[21]の行使のことである。この機構を用いると、もし適切な切断ルールを規定することができれば、協調の進化が促進されることがすでに分かっている。しかし、強ジレンマ領域となると話は別で、いくらネットワークと戦略の共進化による強力な集団ダイナミクスを考慮したとしても、well-mixed 条件下において D 支配が圧倒的な強ジレンマ領域において協調を維持することは容易ではない。これは D エージェントの Temptation (T) が大きくなると同時に、C エージェントが D エージェントと相対した時に食られる影響 (S ; Sucker による) が大きくなることで、D エージェントが得られうる利得が C エージェントに比べて相対的に大きくなり、系を支配する上で圧倒的優位に立つからである。従って、協調の進化を達成するには、協調戦略者の利得減少を如何にして小さくするかが重要なポイントの一つになる。Assortative attachment では、実数値戦略定義を利用し、パラメーター α を導入することで、協調型戦略者同士の緩やかなクラスタリングを効果的に促進し、D エージェントからの exploitation を排除することによって頑強な協調型クラスターを形成し、ジレンマがある領域でも集団を固定ネットワーク上よりも優れた協調的フェーズに保つ事ができる。

図 7.4.2 に示す，均衡時の戦略値分布をみってみる．まず，図 7.4.2(a)に示す弱ジレンマの場合に注目する． w 小の時，エージェントの均衡時の戦略分布幅は $s_i = 0.8 - 1.0$ 付近にまとまっているが， w 大の時，中途半端な協調戦略を保持するエージェントの存在比が増加し，均衡戦略値の分布幅は徐々に大きくなる．図 7.4.2(b)に示す強ジレンマの場合を見てみると， w が小さい場合は裏切り相に転落し，ほぼすべてのエージェントが Nash 均衡戦略である $s=0$ 付近の戦略をとるようになる．一方， w が大きい場合，完全に近い協調戦略をとるエージェントの割合は非常に少ないものの，中途半端に協調する戦略が適度に生き残り，裏切りの戦略と共に協調的戦略が共存できる．このように， w が大きければ強ジレンマ域においてもある程度の協調的戦略が生き残れるのは，**assortative attachment** によるネットワーク進化を許容することで，混合戦略値に基づいた同類のグルーピングが促進され，裏切りを散発的にするような **C-type** エージェント同士で強固なクラスターを形成し高頻度で裏切る **D 型** エージェントからの搾取を緩和させることができるからである．同類のグルーピングが促進されることによって強ジレンマでも **C 型** クラスターが生存する可能性がある一方で，逆に弱ジレンマ領域において完全な協調戦略による支配構造とはならず中途半端な協調戦略を保持するエージェントの生存を許してしまう．これは，ネットワーク進化が速いゆえの弊害で，**D 型** エージェントのクラスタリングも同時に促進されることで戦略の伝播力が弱くなり，それぞれの戦略値群ごとに安定となって，協調型エージェントと中途半端な協調型エージェントの併存平衡に陥ってしまうからである．つまり，**Assortative attachment** は混合戦略の特性をうまく利用し緩やかなグルーピングを促進することで協調の維持を達成するものの，裏切り型戦略を弱ジレンマ域で許してしまうという逆効果もある．

(a) $r = 0.45$



(b) $r = 0.75$

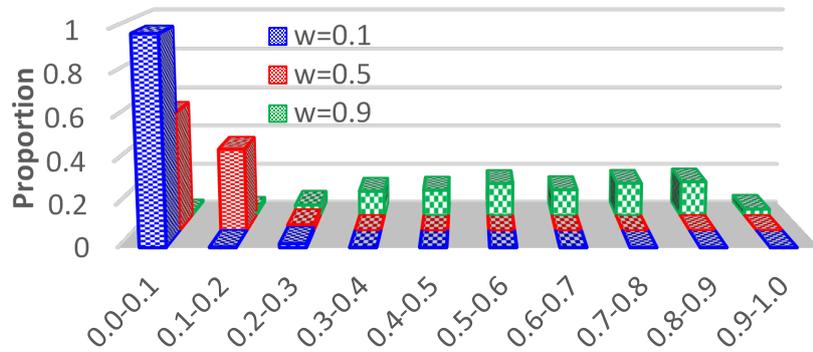


図 7.4.2 (a) $r = 0.45$ または(b) $r = 0.75$ における均衡時の戦略分布. 許容度パラメーター $\alpha = 0.1$. コピーエラーは考慮されていない. ジレンマが強くなると特に w が大きい場合, 戦略分布は広がりを見せる.

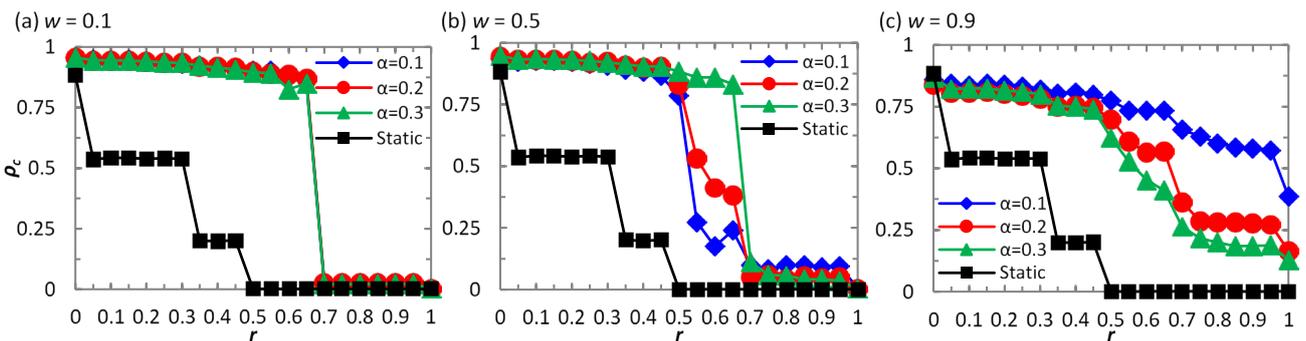


図 7.4.3 (a) $w = 0.1$, (b) $w = 0.5$ (c) $w = 0.9$ それぞれの場合における均衡協調率 ρ_c . "Static" は $w = 0$ すなわち戦略適応のみ生起する状況の結果を表している. 戦略コピーエラーを $\mu = 0.005$ の確率で考慮している

次に、図 7.4.3 に示す、戦略適応プロセスにおいてコピーエラーを加味することによる影響を見てみよう。コピーエラーを加味した場合、固定ネットワーク上のゲームでは均衡にエラー無しとの大きな違いは観察されない。これは、戦略定義に既に確率性を含んだ混合戦略が用いられているからである。一方、提案モデルでは、エラーを加味することで協調がエンハンスされる。

図 7.4.4 に $r = 0.55$ のときの各コピーエラーレートにおける均衡協調率を抜き出して示している。特に w 小の時、コピーエラーを加えすぎない程度の最適な頻度で生起させることで協調が劇的にエンハンスされる。 w が大きいときは、コピーエラーを加味することによるエンハンス効果は見られなくなり、寧ろ協調率を減少させる。エラーを加えたとき、一体エージェントにどのような変化が生じているのだろうか。 w が小さい場合におけるノイズの有無による均衡トポロジーの違いを見てみよう。図 7.4.5 に示す $r = 0.55, w = 0.1$ における均衡時の次数分布をみると、コピーエラーを加味して協調がエンハンスされる場合、エラーを加味しない場合に比べて最高次数が増し、ネットワークのヘテロ性が増していることが分かる。明らかにコピーエラーとヘテロネスには何かしらの関連性がある。

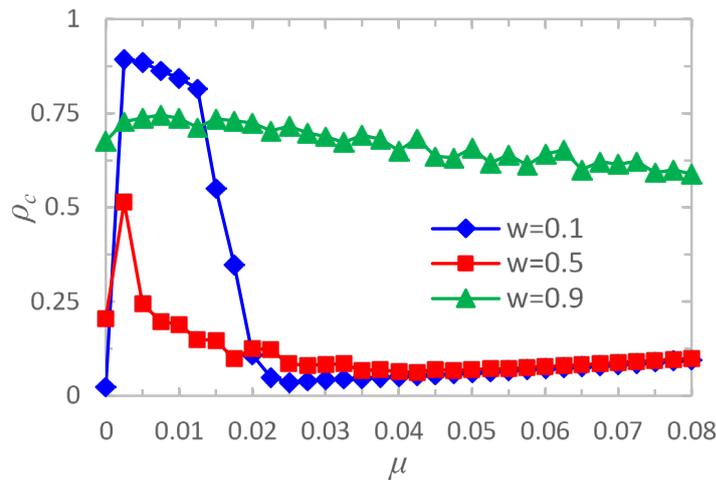


図 7.4.4 コピーエラー率 μ と均衡協調率 ρ_c の関係. $\alpha = 0.1$ の場合を示す.

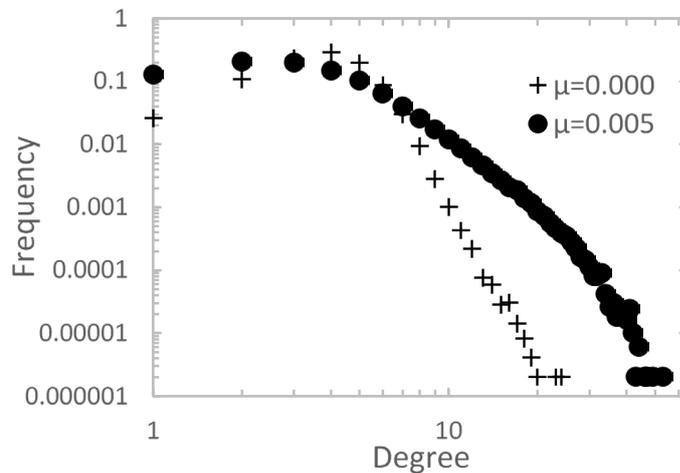


図 7.4.5 $r = 0.55$ における次数分布コピーエラー率 μ と均衡協調率 ρ_c の関係. $w = 0.1$, $\alpha = 0.1$ の場合を示す. エラーが加味されると明らかなトポロジーの差異が観察される.

ヘテロ性増大の理由について検討するために、戦略適応イベントに内在する要素を整理していこう。コピーエラーは、戦略適応においてエージェントが戦略をコピーする際に生起するため、戦略コピーイベントに注目することで系にどのような変化が生じているのかを究明できる。本シミュレーションでは適応方法として IM を用いているため、高利得を獲得しているエージェントは専らコピーされる側となる。ここでいう高利得を獲得するエージェントは、次数が大きなエージェントか、C 戦略に対して D 戦略で対抗した裏切者エージェントである。こういった高利得を獲得するエージェントは自らコピーを行いにくいいため、コピーエラーの影響を大きくは受けない。コピーエラーの影響を受けるのは、こういった高利得者と繋がっている末端の相対的低利得エージェント達である。彼らは自分よりも利得の高いハブの戦略をコピーしながら安定なクラスターの一端を形成していくが、コピーエラーによってハブエージェントとの許容値 α を超える戦略値になるケースが生じる。この場合、この末端

エージェントは、その後のネットワーク適応により相手であるハブエージェントから切断されるかもしくは自ら切断することになる。前者、すなわち高次数 C 型ハブが戦略変位した末端エージェントとのネットワークを切断した場合、そのハブは新たなエージェントとランダムに再接続する。ハブが再接続した相手が D 型のハブではない限り基本的には高頻度で R を取り合っている C 型ハブの方が利得が高いため、新たな隣人エージェントに自戦略をコピーさせて安定な C クラスタを拡大できる。つまり、コピーエラーを加味することで、C 型ハブの安定性を損なわずにリンクの再接続イベントが生起し、それによって末端エージェントのネットワーク進化が促進され、C 型ハブが成長する可能性がある。このような効果によって、エラーを加味することで C 型ハブの切断権行使イベントが増え、結果的にこの C 型ハブの次数がより成長し、ネットワークのヘテロネスが増大、その起結として協調率が改善しているのではないかと考えられる。

ここで、図 7.4.6 に示す w 小のときのジレンマごとの均衡トポロジーを見てみよう。

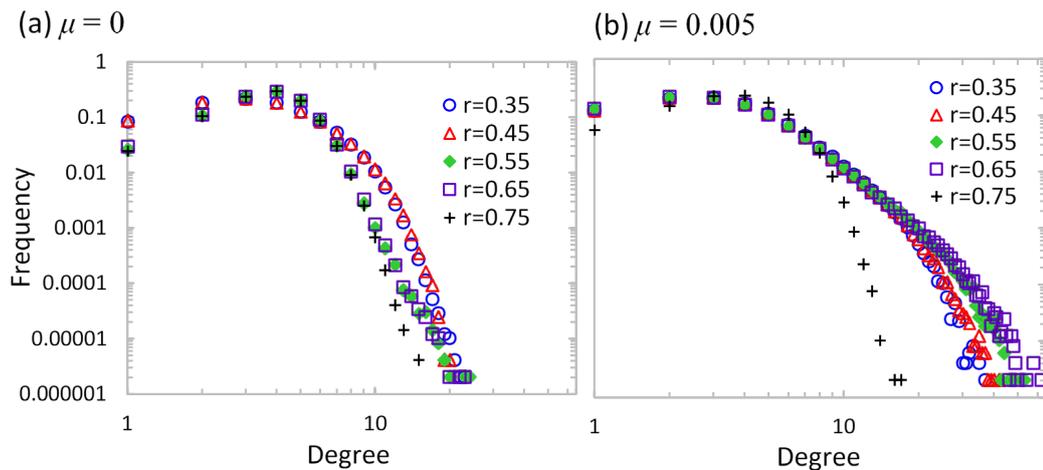


図 7.4.6 $w = 0.1$, $\alpha = 0.1$ のときのジレンマごとの次数分布。臨界ジレンマ付近でヘテロ性が増大する。特にコピーエラーを加えたときほどその傾向は顕著である。臨界ジレンマ付近を超えるとネットワークは進化しない。

ジレンマが弱い部分ではネットワークのヘテロ性はそこまで強くないが、ジレンマが強くなると、ネットワークのヘテロ性が強くなる。その後、協調を保てる臨界ジレンマ強さを超えると、ヘテロ性は再び影をひそめる。これは、ジレンマが強烈すぎるため、ネットワークが進化していく間も無く裏切りフェーズに吸引され、戦略とネットワークの進化が停止してしまうからである。図 7.4.7 に示すように、 w 大が大きくなると、ネットワークにはある程度のヘテロ性があるものの、最高次数はそこまで大きくなく、ヘテロネスが進化しない。つまり、上

記したような C 型ハブのヘテロネスが増大するためには、ネットワーク適応が速すぎることは逆効果となる。

図 7.4.2 で示した戦略分布によると w が大きくなると戦略の分布幅が大きくなっており、その結果、図 7.4.1 にみるように強ジレンマ域では w 大ほど協調がエンハンスされていた。戦略が分布しているということは C 型から D 型が空間的に連続して分居していることを示唆している。従って、 w が大きいときは明確なハブエージェントのようなものは出現せず、D 型エージェントとともに C 型エージェントが生存する無段階的併存平衡的状况が出現し、全体としてみたとき、強いジレンマでもそれなりに協調が維持されるような状況になっていることを示唆している。

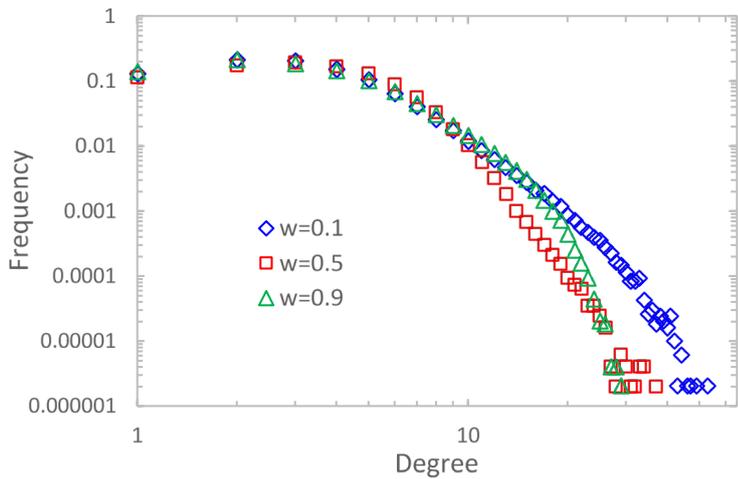


図 7.4.7 ネットワーク適応速度 w と均衡トポロジーの関係. $r = 0.55$, $\mu = 0.005$.

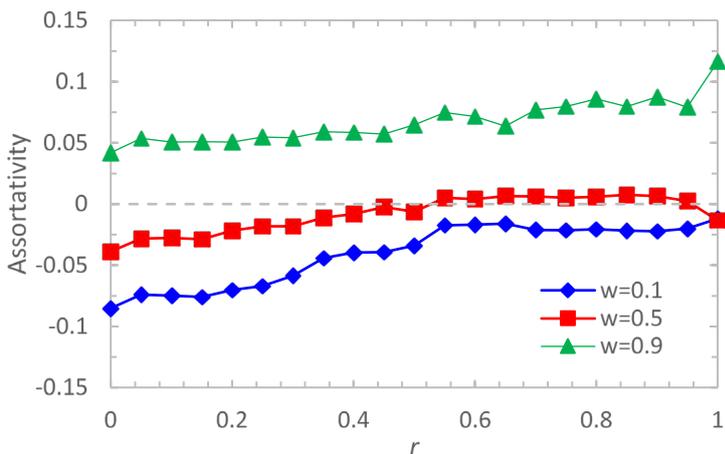


図 7.4.8 $\alpha = 0.1$, $\mu = 0.005$ のときの次数相関. w が大きいほど次数相関が大きくなる。

図 7.4.8 に示す次数相関を見ると、 $w = 0.9$ のときに正の次数相関となっている。高次数ハブが存在しない場合、次数相関が正になるという事は、低次数エージェント同士がそれぞれ小

さなグループに分かれてクラスタリングしているということである。反対に、 w 小の時、次数相関は負値を示しており、比較的高次数の安定なハブに末端のエージェントが取り付いている状況を示唆している。図 7.4.9 が示す、各エージェントの **interaction** における次数関係を見てみると、確かに、 $w = 0.1$ の時、低次数エージェントは平均的に高次数エージェントと繋がりを持っている一方、 $w = 0.9$ の時、高次数エージェントが相対的に減少し、それに伴って低次数エージェント同士の **interaction** の頻度が高くなっている。

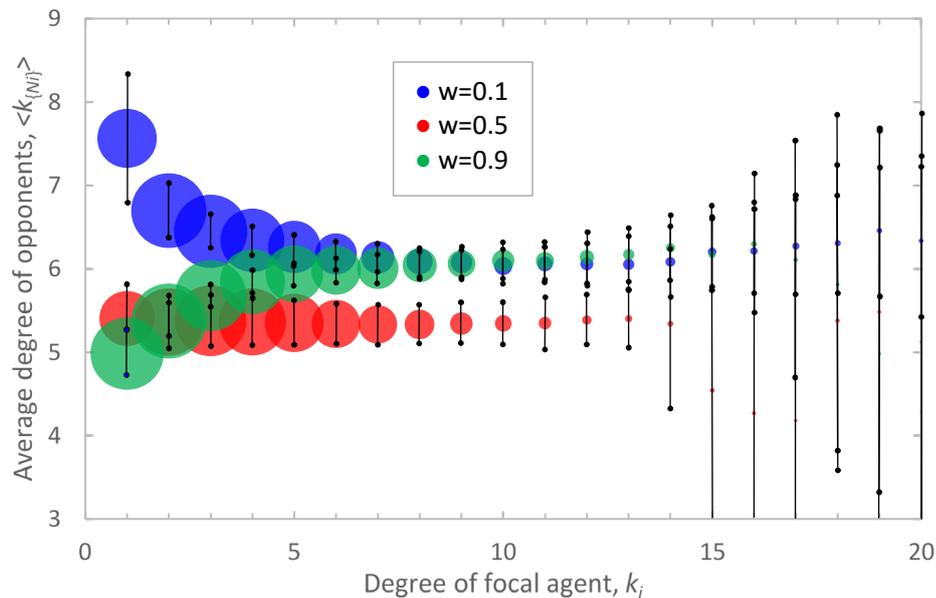


図 7.4.9 各エージェントの次数とその隣人の平均次数の関係. $r = 0.1$ の結果を示す. 黒線は標準偏差を表す. バルーンの大きさは次数 i を保有するエージェントの数を表す.

以上の考察から、 w の大きさによって協調進化のプロセスは異なることが示された。 w が小さい場合、臨界ジレンマ強さ付近では、ヘテロネスの増加が協調の維持に効果的な要素であると考えられ、絶妙な頻度で生起するコピーエラー、ネットワーク適応速度と戦略適応速度のバランスによって安定なハブが成長し、ヘテロネスを増大させることで固定ネットワーク上よりもさらに協調の進化を達成する。 w 大の時は、安定なハブに頼らず（協調度の高いハブに率いられるクラスターは成長せず）、そこそこ協調する C 型戦略を持った低次数クラスターを中心に、類似戦略を保有するクラスターと連続的にクラスターを形成することで裏切りの戦略からの侵襲を防いでいる。次数を大きくして同様のシミュレーションを行うと、協調が減衰するが、依然固定ネットワークよりも成績は良く、**tolerance parameter** やノイズに対する応答は同一のものであった。

7.5. 結論

混合戦略を前提とし、保有する戦略と同類の戦略を持つエージェント同士とクラスタリングするような *Assortative Attachment rule* というネットワーク進化プロトコルを仮定した共進化モデルを提示した。この枠組みを用いることで、固定ネットワーク上でのゲームよりもより協調戦略が進化することが分かった。従来のネットワーク適応ルールは利得や戦略の違いでネットワークの保持、切断を決めるものが多かった。これらの研究は、人間を含む生物が潜在的に高利得を欲したり、自分への有利不利で繋がりを解消したりする行動が、協調を強力に進化する理由の一つであることを構成論的に示すものである。このような視点からさらに踏み込み、本枠組みでは、人や生物の、“同種を好む傾向”をモデルに組み込むことで、同類的なエージェントの自己組織化が協調の維持に有効であることを示した。このルール、*Assortative Attachment* では、どのエージェントに関しても、保有戦略値に応じて公平にリンクを再構成する機会が与えられるため、Zimmermann を初めとする既往の共進化モデルにおける C エージェントの生き残りや、D エージェントの生き残りに対しての恣意性が全く仮定されていない。すなわち、協調者にとっても裏切者にとっても非常に公平なネットワーク適応ルールであると言えるだろう。この状況下において協調創発の可能性を示せたという事は、人や生物が *real interactions* において協調行動を維持できるのは、同類性を志向するからであると云えるかもしれない。また、本枠組みにおいて、 w 小の時に、ネットワークのヘテロネスと協調の進化に関連性が見出された。ネットワークの多様性ともいえるヘテロネスは、協調進化におけるキーファクターとして多くの研究が行われているが[112]、本研究もまた、ネットワークのヘテロ性が協調の進化に関連しているということを示唆する結果となった。

8. 総括結論

8.1. 結論

本論では、進化ゲーム理論に基づく数理モデルを構築し、個人の利益を最大化する行動と社会全体にとって最適な選択とが異なるジレンマ環境下における協調の創発メカニズムに関する一連の研究成果を報告した。

第 1 章においては古典ゲーム理論および進化ゲーム理論の成り立ちおよび既往の研究についての包括的なレビューを通し、協調創発の 5 つの互惠機構の一つであるネットワーク互惠を組み込んだ進化ゲームモデルにおいて、動的ネットワークおよび混合戦略モデルの組み合わせにおける更なる研究の必要性を示した。また、複数のジレンマクラスにおいてもユニバーサルに適用できるような協調創発のメカニズムについてより深い追求が必要である点を示し、本論における研究の方向性及び意義を見出した。

第 2 章では、本論の研究において中核となる、 2×2 (2 人 2 戦略) ゲームおよび **Ultimatum Game** (最後通牒ゲーム) の基礎理論について詳述した。本論では進化ゲームに基盤ネットワークを仮定したネットワークゲームが用いられたため、複雑ネットワークにおける基礎パラメーターおよびネットワークの成り立ち、ネットワークにて定義されたエージェント同士の進化のための戦略適応方法について述べた。

第 3 章では、**Ultimatum Game** において相手へのオファーが増幅されるメカニズムを考慮した拡張 **Ultimatum Game** を定義し、実数 2 次元の戦略であるオファーとオファー許容値の適応に加え、リンク切断によるネットワーク適応を考慮した共進化モデルを構築した。様々なリンク切断プロトコルを提示した数値実験を通し、進化の帰結は共進化の切断プロトコルにより大きく異なることが明らかとなった。切断プロトコルによっては社会的に好ましい帰結をもたらすものがあるが、切断プロトコル中に、社会科学的に見て不自然な“恣意性”があってはならない。本研究から、インセンティブ基準切断プロトコルは自他のオファーと許容閾値の大小関係に応じてリンク切断権付与する、社会的には「理に適った」プロトコルであり、実証実験とも整合性のとれる結果となることを導いた。

第 4 章では、 2×2 ネットワークゲームにおいて、対戦相手の前回の戦略に応じて異なる戦略を自身の保有する実数戦略値に応じた確率でとることを許容した一記憶長混合戦略モデルを仮定した。エージェントに記憶長 (メモリー) を考慮することは、協調創発のための代表的な 5 つの互惠機構の中の直接互惠に該当する。本章では、ネットワーク互惠と直接互惠とい

った異なる二つの互惠機構を進化ゲームに考慮することによって、囚人のジレンマとチキンゲームといった異なるジレンマクラスにおいても、各ジレンマクラスにおいてそれぞれ最適な社会帰結である R 互惠や ST 互惠を創発させるような最適戦略をエージェントが可塑的に選択するように進化していくことが分かった。社会に内在するジレンマクラスは一つとは限らないため、複数のジレンマクラス下において各々最適な戦略を自律的に選択していくように進化する機構は現実社会における生物の振る舞いに近いものであるということができよう。本章における発見は、実社会における協調の創発メカニズム解明・数理モデル化という観点で意義のあるものだったと考えられる。

第 5 章では、自身の保有する実数戦略値に応じた確率で協調か裏切戦略をとることを決定する混合戦略モデルにおける協調行動創発機構を詳しく検証した。既往研究から、混合戦略モデルに内包される確率的挙動によって裏切戦略をとるエージェントの一人勝ちを防ぐことで協調クラスターの拡大が可能となることは分かっていた。本章では、混合戦略モデルそのものの戦略更新ルールを厳密に定義し、戦略を *Synchronous* に更新するか、*Asynchronous* に更新するかという違いにまで立ち返り数理モデルの成り立ちを分析した。検証の結果、混合戦略モデルにおける確率的揺らぎによる共鳴現象によって協調クラスターが偶発的に拡大し、強ジレンマ域及び社会粘性の低い高ネットワーク次数環境下においても協調戦略の生き残りが可能なシナリオが存在するというを示した。これは、混合戦略モデルが協調の創発に寄与するメカニズムをより詳細に解き明かしたものであり、進化ゲームモデルにおける数理的な挙動と実社会における協調創発メカニズムの解明という両面において大きな成果であったといえる。

第 6 章では、第 5 章に引き続き混合戦略モデルにおける協調の創発メカニズムについて更に研究するため、既存の離散戦略モデルと完全な混合戦略モデルとを連続的に定義できるようにし、離散戦略から中間的な混合戦略、完全な混合戦略という具合に各戦略定義による進化の帰結の違いを分析した。本章における研究により、完全な混合戦略ではなく、一部でも確率的に戦略を決定する混合戦略的エージェントが系に存在することによって、協調クラスターの生存が可能となることを発見した。この発見は第 5 章によって導き出された混合戦略による確率共鳴と協調の進化の関係性をより強固に支持する結果であったといえる。また、もたらされた結果を見ると、どっちつかずの日和見戦略をとるエージェントの存在が協調の進化に寄与していると言えることから、社会的にも示唆に富んだ結果を得ることができたといえるだろう。

第7章では、混合戦略モデルにおいて動的ネットワークを仮定した。混合戦略モデルに対する戦略とネットワークの共進化の研究は多くは行われていなかった。本章では、エージェントの持ちうる戦略値の相似性（同類性）に着目したネットワーク適応プロトコルを提示した。既存の離散戦略モデルで用いられているネットワーク適応プロトコルは、相手の戦略や利得に基づいて確定的に切断・再接続を行うものとなっていることから、協調の進化や社会にとって好ましい方向の進化を意図的に促進する機構であると言えなくもなかった。本章で提示されたプロトコルを用いることによって、エージェントの戦略的同類性というニュートラルな概念に基づく切断ルールを用いたとしても協調の創発が自発的に促されることを明らかにした。本切断ルールは混合戦略モデルでのみ定義できるルールであり、新たな進化ゲーム研究の視点を提示することができたと考える。また、同類性という社会にとって普遍的な原理を進化ゲーム理論モデルに仮定することで協調の進化を観察できたという点は、社会的にも示唆に富む結果を得られたと考える。

8.2. 今後の展望

本論でのモデル基盤理論として援用した進化ゲーム理論ではモデル中の意思決定主体；エージェントの行動様式を協調または裏切りの2種類に分割することで人と人（生物と生物）との間で生じる相互作用を簡潔に記述することができる。この様式で相互作用を記述することで、エージェント間の相互作用をボトムアップ的に再現し、結果的に複雑現象の観察に至る。この手法をマルチエージェントシミュレーション (Multi Agent Simulation ; MAS) と呼ぶ。この手法の利点は、一般向けワークステーションレベルの計算能力でも十分に進化計算を行える点にある。また、エージェント間の相互作用や行動原理をゲーム理論に基づき大胆に簡略化しているため、こまかな条件をモデル中に再現していない。シンプルなルールで複雑現象を再現できるのがマルチエージェントの大きな長所である。進化ゲーム理論 MAS モデルを用いれば、古典的ゲーム理論では解けなかった（＝社会全体として好ましい状態に均衡させることができなかった）社会ジレンマ構造化においても社会全体の効用が高まるような方向に系が進化するという帰結を得られる。進化ゲーム理論研究におけるシミュレーションモデルがある程度現実の社会を表す数理モデルとしての側面を持っているとすれば、社会ジレンマが解かれるメカニズムを分析することで複雑系社会における現実の社会ジレンマを解消するメカニズムを分析する一助になりえるだろう。

近年、交通流のシミュレーションモデルにおいて、進化ゲーム理論の概念を用いて交通流

動モデルの分析を行い、ドライバーの戦略の進化や、付与される戦略によって社会の全体利得が高まる方向に進化する状況を解明する試みが行われている。また、進化ゲーム理論の枠組みにワクチン接種の概念を組み込むことで、人間社会におけるウイルスの伝搬や、ウイルスの伝搬を防ぐためにワクチン接種を促進するための最適な戦略や社会システムの状態を示唆するという進化ゲーム理論の応用研究も盛んになってきている。現実の現象を数理モデル中により忠実に再現しようとする、必然的に現象を再現するためのパラメーターを多数導入する必要がある。発展モデルにおいて現象を考察する時、数理シミュレーションモデル中に仮定する数式・パラメーターによって結果が大きく変わってくる場合があり、これは仮定する数字によってシミュレーションを自分の意図した方向にコントロールできてしまう恣意性があるともいえるが、仮定する数字によって変遷するシミュレーション結果を正しく分析することで進化のダイナミクスや進化ゲーム理論ベースの数理モデルに対する深い考察を得ることができるともいえる。また、応用モデルを駆使する際、ベースとなる進化ゲーム理論モデルそれ自体の成り立ちが疑わしいものであれば結果そのものは意味をなさない。なぜなら基盤モデルそれ自体は人間の相互作用をそれらしく模擬していることが前提にあるからである。つまり、使用する基盤モデルには十分信頼に足る、確証づけられたものを用いなければ、シミュレーションを用いた帰納的アプローチによって有用な知見を得られているとは言えないだろう。応用研究として現実の複雑系社会における現象に進化ゲーム理論における既往の枠組みを適用した応用モデルを構築し適切に結果を分析するためには、基礎研究としての進化ゲーム理論とその数理モデルそのものに対する研究はやはり必要不可欠であるといえるだろう。

本論では原型の進化ゲームモデルをベースに、ネットワーク構造や動的ネットワーク、直接互惠といった基礎的な既往の研究理論を組み合わせ、主に混合戦略モデルによる確率的要素がどの様に相互作用を及ぼすのかという視点を導入し、協調行動の進化の様子を検証した。離散戦略モデルや進化ゲームの根幹を成す互惠機構そのものに対する研究が進んできた今、生物・社会のランダムな振る舞いを表現するという観点から、混合戦略モデルに内包されるような確率的要素や複数の互惠機構との組み合わせや動的ネットワーク等を進化ゲームモデルに導入し、適切に分析していくことは進化ゲームモデルのさらなる発展に重要であると考えられる。本論によって提示されたモデルは、既往研究によって得られた知見を統合し発展させたものであるため、本論にて提示された進化ゲームモデルによってより多様な社会的相互関係を進化ゲームモデル上で数理的に表現できるようになったといえるだろう。ここで、既往

の進化ゲームモデル研究を基礎研究とするならば、ワクチン接種モデルや交通流モデルは進化ゲームの枠組みを応用した発展モデルといえるだろう。本論による新たな発見によってより複雑な相互作用を進化ゲームで表現できるようになったことから、本論の成果が基礎研究の枠組みにとらわれず、発展モデルへの展開という観点でも進化ゲーム理論適用の可能性を広げたい。さらに近年では、実際のビジネスの範疇においても、従来の進化ゲームの枠組みで用いられてきていた戦略適応機構の一つである遺伝的アルゴリズムを用い、複数の要素が組み合わされた複雑な状況における最適解探索を行うということも一般的になりつつあり、人工知能分野における学習アルゴリズムへの適用という側面でも進化ゲームの知見応用の可能性は広がっているとと言えるだろう。本論が多少なりとも今後の進化ゲーム理論研究の発展に寄与することを祈念し、総括とする。

参考文献

- [1] Smith J.M., *Evolution and the Theory of Games* (Cambridge University Press, 1982).
- [2] Nowak M.A., *Science* 314, 1560 2006.
- [3] Yamauchi A., Tanimoto J. and Hagishima A., *BioSystems* 103, 85–92 2011.
- [4] Shigaki K., Tanimoto J., Wang Z., Kokubo S., Hagishima A. and Ikegaya N., *Phys. Rev. E* 86, 031141 2012.
- [5] Szabó G. and Tóke, C., *Phys. Rev. E* 58, 69–73, 1998.
- [6] Zimmermann M.G., Eguiluz V.M. and San Miguel M., *Phys. Rev. E* 69, 065102R, 2004.
- [7] Tanimoto J., *Physica A* 388, 953-960, 2009.
- [8] Güth W. and Tietze R.; *Ultimatum Bargaining behavior, a survey and comparison of experimental result*, c11, 417-449, 1990.
- [9] Page K.M., Nowak M.A. and Sigmund K., *Proc R Soc Lond B*, 267:2177, 2000.
- [10] Nowak M.A., Page K.M. and Sigmund K., *Science*, 289:1773, 2000.
- [11] Roth A., Prasnikar V., Okuno-Fujiwara M. and Zamir S., *Am Econ Rev*, 81:1068, 1991.
- [12] Roth A., *Bargaining experiments in handbook of experimental economics*. Princeton University Press; 1995.
- [13] Kishimoto N., Kokubo S. and Tanimoto J., *Chaos Solit. Fractals* 56, 83-90, 2013
- [14] Kokubo S., Wang Z. and Tanimoto J., *Appl. Math. Comput.* 259, 552-568, 2015
- [15] Güth W. and Tietze R., *J Econ Psychol*, 11:417, 1990.
- [16] Thaler R.H., *J Econ Perspect*, 2:195, 1988.
- [17] Fehr E. and Schmidt K., *Q J Econ*, 114:817, 1999.
- [18] Sigmund K. and Nowak M.A., *Sci. Am.*, 286:82, 2002.
- [19] Li X. and Cao L., *Phys. Rev. E*, 80:066101, 2009.
- [20] Iranzo J., Floría L.M., Moreno Y. and Sánchez A., *PLoS ONE*, 7:e43781, 2012.
- [21] Schuessler R., *J Conflict Resolut*, 33:728, 1989.
- [22] Page K.M. and Nowak M.A., *Bull Math Biol*, 64:1101, 2002.
- [23] Sánchez A. and Cuesta J.A., *J. Theor. Biol.*, 235:133, 2005.
- [24] Duan W.-Q. and Stanley H.E., *Phys. Rev. E*, 81:026104, 2010.
- [25] Da Silva R. and Kellerman G.A., *Braz J. Phys.*, 37:1206, 2007.
- [26] Da Silva R., Kellerman G.A. and Lamb L.C., *J. Theor. Biol.*, 258:208, 2009.
- [27] Killingback T. and Studer E., *Proc. R. Soc. London B*, 268:1797, 2001.
- [28] Iranzo J., Román J. and Sánchez A., *J. Theor. Biol.*, 278:1, 2011.

- [29] Sinatra R., Iranzo J., Gómez-Gardeñes J., Flori' a L.M., Latora V. and Moreno Y., *J. Stat. Mech.*, 9:P09012, 2009.
- [30] Szolnoki A., Perc M. and Szabó G., *EPL*, 100:28005, 2012.
- [31] Lieberman E., Hauert C. and Nowak M.A., *Nature*, 433:312, 2005.
- [32] Deng L.-L., Tang W.-S. and Zhang J.-X., *Chin. Phys. Lett.*, 28:080204, 2011.
- [33] Deng L.-L., Tang W.-S. and Zhang J.-X., *Physica A*, 390:4227, 2011.
- [34] Gao J., Li Z., Wu T. and Wang L. L., *EPL* 93:48002, 2011.
- [35] de Jong S., Uyttendaele S. and Tuyls K., *J Artif. Intell. Res.*, 33:551, 2008.
- [36] Szabo G. and Fath G., *Phys. Rep.* 446: 97–216, 2007.
- [37] Perc M. and Szolnoki A., *BioSystems* 99: 209–125, 2010.
- [38] Tanimoto J., *Sociobiology* 58: 1–11, 2011.
- [39] Kim B.J., Trusina A., Holme P., Minnhagen P., Chung J.S., et al., *Phys Rev E* 66: 021907, 2002.
- [40] Vainstein M.H. and Arenzon J.J., *Phys. Rev. E* 64: 051905, 2001.
- [41] Wang Z., Wang Z., Zhu X. and Arenzon J.J., *Phys. Rev. E* 85: 011149, 2012.
- [42] Traulsen A., Pacheco J.M. and Nowak M.A., *J. Theor. Biol.* 246: 522–529, 2007.
- [43] Roca C.P., Cuesta J.A. and Sa' nchez A., *Europhys. Lett.* 87: 48005, 2009.
- [44] Helbing D., Szolnoki A., Perc M. and Szabo' G., *PLoS Comput. Biol.* 6: e1000758, 2010.
- [45] Santos F.C., Pacheco J.M. and Lenaerts T., *PLoS Comput. Biol.* 2: e140, 2006.
- [46] Xia C.Y., Wang J., Wang L., Sun S.W. and Wang J.S., *Chaos Solt. Fractals* 45: 1239–1245, 2012.
- [47] Poncela J., Gomez-Gardenes J., Traulsen A. and Pacheco J.M., *New J. Phys.* 11: 083031, 2009.
- [48] Helbing D., Szolnoki A., Perc M. and Szabo' G., *New J. Phys.* 12: 083005, 2010.
- [49] Pierre B. and Marco T., *Phys. Rev. E* 86: 066107, 2012.
- [50] Tang C.B., Xiang L., Lang C. and Zhang J.Y., *J. Theor. Biol.* 306: 1–6, 2012.
- [51] Zhang H.F., Yang H.X., Du W.B., Wang B.H. and Cao X.B., *Physica A* 389: 1099–1104, 2010.
- [52] Hauert C. and Doebeli M., *Nature* 428: 643–646, 2004.
- [53] Tanimoto J. and Sagara H., *BioSystems* 90: 105–114, 2007.
- [54] Tanimoto J. and Sagara H., *BioSystems* 90: 728–737, 2007.
- [55] Szolnoki A., Xie N.G., Wang C. and Perc M., *Europhys. Lett.* 96: 38002, 2011.
- [56] Xia C.Y., Ma Z., Wang Y., Wang J. and Chen Z.Q., *Phys. Scripta* 84: 025802, 2011.
- [57] Zhong W., Kokubo S. and Tanimoto J., *BioSystems* 107, 88
- [58] Tarnita C.E., Antal T. and Nowak M.A., *J. Theor. Biol.* 261, 50, 2009.

- [59] Chen C., Fu F., and Wang L., *Phys. Rev. E* 78, 051120, 2008.
- [60] Santos F.C., Pinherio F.L., Lemaerts T. and Pacheco J.M., *J. Theor. Biol.* 299, 88, 2012.
- [61] Moyano L.G. and S´anchez A., *J. Theor. Biol.* 259, 84 , 2009.
- [62] Roca C.P., Cuesta J.A., and S´anchez A., *Phys. Rev. E* 80, 046106 , 2009.
- [63] Grilo C. and Correia L., *J. Theor. Biol.* 269, 109 , 2012.
- [64] Antonioni A. and Tomassini M., *Advs. Complex Syst.* 15, 1250046 , 2012.
- [65] Zimmermann M.G., Eguíluz V.M. and Miguel M.S., *Phys. Rev. E*, 69:065102, 2004 .
- [66] Santos F.C. , Pacheco J.M. and Lenaerts T., *PLoS Comput Biol*, 2:e140, 2006.
- [67] Poncela J. , Gómez-Gardeñes J. , Floría L.M. , Moreno Y. , Sánchez A. . *Europhys. Lett.*, 88:380 03, 2009.
- [68] Lee S., Holme P., Wu Z.X., *Phys. Rev. Lett.*, 106:028702, 2011.
- [69] Pacheco J.M., Traulsen A., Ohtsuki H. and Nowak M.A., *J. Theor. Biol.*, 250:723, 2008.
- [70] Szolnoki A. and Perc M, *Europhys. Lett.*, 86:30007, 2009.
- [71] Moreira J.A., Pacheco J.M. and Santos F.C., *Sci. Rep.*, 1521, 2013.
- [72] Cardillo A., Meloni ., Gómez-Gardeñes J. and Moreno Y., *Phys. Rev. E*, 85:067101, 2012.
- [73] Wang Z., Szolnoki A. and Perc M., *Sci. Rep.*, 2:369, 2012.
- [74] Wang Z, Szolnoki A. and Perc M., *Phys. Rev. E*, 85:037101, 2012.
- [75] Sicardi E.A., Fort H., Vainstein M.H. and Arenzon J.J., *J. Theor. Biol.*, 256:240, 2009.
- [76] Ichinose G., Saito M., Sayama H., and Wilson D.S., *Sci. Rep.*, 3:2509, 2013.
- [77] Szolnoki A., Perc M. and Danku Z., *EPL Europhys. Lett.*, 84:50007, 2008.
- [78] Yang Z., Li Z., Wu T. and Wang L., *Phys, A Stat. Mech. Its Appl.*, 392:2038, 2013.
- [79] Poncela J., Gómez-Gardeñes J., Traulsen A. and Moreno Y., *New J. Phys.*, 11:083031, 2009.
- [80] Liu J., Meng H., Wang W., Xie Z. and Yu Q., *Appl. Math. Comput.*, 340:234, 2019.
- [81] Jia D., Shen C., Li X., Boccaletti S. and Wang Z., *EPL*, 127:68002, 2019.
- [82] Chu C., Hu X., Shen C., Li T., Boccaletti S., Shi L. and Wang Z., *Chaos Solt. Fractals*, 29:013139, 2019.
- [83] Nowak M.A., *Evolutionary Dynamics: Exploring the Equations of Life* (Belknap Press, n.d.).
- [84] Masuda N., 複雑ネットワークの科学 (産業図書, 2005).
- [85] Erdős A.R.P., *On the Evolution of Random Graphs* (PUBLICATION OF THE MATHEMATICAL INSTITUTE OF THE HUNGARIAN ACADEMY OF SCIENCES, 1960), pp. 17–61.
- [86] Watts D.J. and Strogatz S.H., *Nature* 393, 440, 1998.
- [87] Newman M., *Phys. Rev. Lett.* 89, 208701, 2002.

- [88] Barabási A., *Science* (80-.). 286, 509, 1999.
- [89] Yamauchi A., Tanimoto J. and Hagishima A., *Biosystems*. 103, 85, 2011.
- [90] Deng, L.-L., Tang, W.-S. and Zhang, J.-X., 390, 4227-4235, 2011.
- [91] Nowak M.A. and Sigmund, K, *Acta Applicandae Math.* 20, 247-265, 1990.
- [92] Metz, J.A.J., Geritz, S.A.H., Meszena, F.J., Jacobs, F.J.A. and van Heerwaardenm, J.S., (Van Strien, S.J., Verduyn Lunel, S.M., eds), 183-231, Amsterdam, North Holland, 1996.
- [93] Gao J., Li Z., Wu T. and Wang, *EPL* 93, 48002, 2011.
- [94] Tanimoto J., *Int. J. Mod. Phys.C22* (4), 401-407, 2011.
- [95] Santos F.C., Pacheco J.M. and Lenaerts T., *PLoS Comput. Biol.* 2 (10), 1284–1291, 2006.
- [96] Tharler R.H., *J. Econ. Pers.* 2, 195-206, 1988.
- [97] Güth W. and Tietze R., *Ultimatum Bargaining behavior, a survey and comparison of experimental result*, c11, 417-449, 1990.
- [98] Roth A.E., *Bargaining experiments*, in *Handbook of Experimental Economics*, Princeton University Press, 1995.
- [99] Fehr E. and Achmind K., *Quart. J. Econ.* 114, 817-868, 1999.
- [100] Sigmund K. and Nowak M.A., *Sci. Am.* 286, 82-87, 2002.
- [101] Bolton G.E. and Zwick R., *Games Econ. Behav.* 10, 95-121, 1995.
- [102] Miyaji K., Tanimoto J., Wang Z., Hagishima A. and Ikegaya N., *PLoS One* 8, e71961 , 2013.
- [103] Tanimoto J. and Sagara H., *Biosystems* 90, 105 , 2007.
- [104] Zhong W., Kokubo S. and Tanimoto J., *Biosystems*. 107, 88 , 2012.
- [105] Wang Z. and Perc M., *Phys. Rev. E* 82, 021115 , 2010.
- [106] Hauert C. and Doebeli M., *Nature* 428, 643 , 2004.
- [107] Miyaji K., Tanimoto J., Hagishima A. and Ikegaya N., *Phys. Rev. E* 88, 062149 , 2013.
- [108] Shigaki K., Wang Z., Tanimoto J. and Fukuda E., *PLoS ONE* 8, e76942, 2013.
- [109] Tanimoto J., *Phys. Rev. E* 87, 062136 , 2013.
- [110] Tanimoto J., *Phys. A Stat. Mech. Its Appl.* 389, 3325, 2010.
- [111] Santos F.C., Pacheco J.M., and Lenaerts T., *PLoS Comput. Biol.* 2, e140, 2006.
- [112] Nowak M., Santos F.C., Pinheiro F.L., Lenaerts T. and Pacheco J.M., *J. Theor. Biol.* 299, 88, 2012.

謝辞

本論文を制作するにあたり指導して下さった谷本潤先生，萩島理先生，池谷直樹先生に心から感謝いたします。