卒業論文 2021 年度 (令和 3 年)

ビザンチン将軍問題の解法に基づいた
計算可能な複雑系としての社会契約

慶應義塾大学 環境情報学部
宮元 眺

卒業論文要旨 - 2021 年度 (令和 3 年度)

ビザンチン将軍問題の解法に基づいた計算可能な複雑系としての社会契約

インターネットのグローバル化や Bitcoin[17] を始めとするブロックチェーン技術の登場により、インターネット上のサービスの正当性を特定の国家の法によって保証することは困難になりつつある。我々は、社会契約の概念をインターネットに適用することがこの問題の解決に繋がると考え、それを計算可能なレベルまで抽象化することを試みている。社会契約とは、ある集団においてその成員達が合意された約束を必ず履行する状態に至るプロセスである。現代においては、この概念は交渉理論や進化ゲーム理論を用い、正義や倫理、道徳といった概念とともに研究がなされている [27] が、その多くは利得の分配を強制執行する力の存在が暗黙的に仮定されている。この強制執行力は、それ自身が利得を分配する約束であるため、集団の外部にその存在を仮定することは別の社会契約の成立を仮定することに等しい。それゆえ、社会契約全体をモデリングするためには、強制執行力が集団の内部で生じるメカニズムを解明する必要がある。Binmore の 2005 年の研究 [5] はこれを試みているが、このアプローチは還元主義的なため強制執行力を生じさせる成員の振る舞いが具体的にどういったものか明確でない。本研究では、強制執行力が集団の外部に存在しない場合に社会契約が成立するのかという問いに取り組む。ここで我々は、強制執行力を自己組織化させる成員の振る舞いがビザンチン将軍問題の署名付きの解法 [15] によって説明でき、集団の一部のみがそのように振る舞う場合であっても社会契約が成立することを示す。そのために我々はゲーム理論とビザンチン将軍問題の解法を用いて社会契約を成員の振る舞いによる複雑系としてモデリングし、ランダムな 8 体のエージェント (8 種類・重複あり) の相互作用の果てに社会契約が成立するか確かめるシミュレーションを行った。64000 回の施行の結果、誠実に振る舞うエージェントが過半数を超える場合、サンプリングした全てのケースで社会契約は成功した。これにより、社会契約の成立に必要となる歴史の定義や社会指標の決定方法、各成員の振る舞いを計算可能なレベルで明確にし、同時に、強制執行力が集団の外部に存在しない場合でも過半数の成員が誠実であれば社会契約が必ず成立することを示した。この誠実な成員と社会契約の成立の関係性はヒューリスティックな解であるため、より厳密な解を得るには大規模な総当り実験を行うか理論的な解明がなされる必要がある。仮に $3f+1$ よりゆるい条件で合意された約束が必ず履行される状態に到れるならば、時間経過とともに f が増加する場合に $3f+1$ 以下の人数でフォールトトレラントを達成できる実用的な BFT アルゴリズムを示せる可能性がある。また、現実の社会を考えたとき、集団の人数は時間経過とともに変化するものであるため、それを考慮して社会契約を更新するモデルの構築が必要である。本研究がそういった新たな研究の足がかりになることを願う。

キーワード:
1. 社会契約, 2. ビザンチン将軍問題, 3. 複雑系, 4. 評判システム

慶應義塾大学 環境情報学部
宮元 眺

卒業論文要旨 - 2021 年度 (令和 3 年度)

Social Contract as a Computable Complex System Based on the Solution of the Byzantine Generals Problem

With the globalization of the Internet and the emergence of blockchain technologies such as Bitcoin[17], it is becoming increasingly difficult to guarantee the legitimacy of services on the Internet through the laws of a particular nation. We believe that applying the concept of social contract to the Internet will help solve this problem, and we are attempting to abstract it to a computable level. The social contract is the process by which a group of people reach a state in which they always fulfill their agreed-upon promises. In modern times, this concept has been studied using bargaining theory and evolutionary game theory, along with concepts such as justice, ethics, and morality[27]. Many of them implicitly assume the existence of the *enforcement* to the distribution of payoff. Since this *enforcement* is itself a promise to distribute payoff, assuming its existence outside the group is equivalent to assuming the formation of another social contract. Therefore, in order to model the social contract as a whole, it is necessary to elucidate the mechanisms by which *enforcement* arises within a group. Binmore's 2005[5] study attempts to solve this problem, but because this approach is reductionist, it is not clear what exactly are the behaviors of members that give rise to *enforcement*. In this study, we research the question of whether a social contract can be established when the *enforcement* does not exist outside the group. Here we show that the behavior of members who self-organize *enforcement* can be explained by the signed solution to the Byzantine Generals Problem[15], and that a social contract can be established even when only a part of the group behaves in this way. For this purpose, we modeled the social contract as a complex system with members' behavior using game theory and the solution of the Byzantine General Problem. We conducted a simulation to see if a social contract can be established after the interaction of eight random agents (eight types, with overlaps). As a result of 64,000 runs, the social contract was successful in all the cases sampled when the majority of agents behaved honestly. In this way, we have been able to define the history required for the formation of a social contract, determine the social indicators, and clarify the behavior of each member on a computable level. At the same time, it showed that a social contract can always be established if the majority of members are sincere, even if there is no *enforcement* outside the group. Since the relationship between sincere members and the formation of a social contract is a heuristic solution, a large-scale brute force experiment or theoretical clarification is needed to obtain a more rigorous solution. If we can reach a state in which promises agreed to under looser conditions than $3f+1$ are always fulfilled, then We may be able to show a practical BFT algorithm that can achieve fault-tolerance with less than $3f+1$ people when f increases with time. In addition, when we consider real society, the number of people in a group changes over time, and it is necessary to construct a model that takes this into account and updates the social contract. I hope that this research will serve as a stepping stone for such new research.

Keywords:
1. Social Contract, 2. Byzantine Generals Problem, 3. Complex System, 4. Reputation System

Keio University Faculty of Environment and Information Studies
Nozomu Miyamoto

目 次

第1章　序論

1.1　本研究の動機

社会契約とは、ある集団においてその成員達が合意された約束を必ず履行する状態に至るプロセスである。中世ヨーロッパにおいて、この概念はホッブズやロック、ルソーによって議論され「国家において、なぜ国民が法に従うのか」という問いに理論的な根拠をもたらすことで王権神授説を否定する主張として注目された。現代においては、ロールズ [18] やハーサニー [13] が中世の社会契約論の一般化を試み、そこから導きだされる公正としての正義について議論した。それ以降、社会契約は正義や倫理、道徳といった概念と結びつけて議論されることになる。社会契約の研究の意義は社会の変化とともに移り変わっており、社会の動向の変化に伴って今後も様々な意義が生まれるだろう。その一つの例として、我々はインターネットのもたらす社会の変化が社会契約の研究に新たな意義を生じさせると考えている。

2020 年の GSMA[11] の調査では、モバイル回線のユニークな契約者数は 5.8 億人に登り、世界人口の 70% がモバイル回線を所有していることが示唆されている。インターネットユーザーが全世界で増加する一方、Google[2] や Facebook[1] のようなサービスは全世界にユーザーを抱えるまで成長し、それらの運営会社も多国籍企業として多数の国に支社を置くようになっている。こうした企業のガバナンスは、もはや特定の国家の法によって管理するのは困難なレベルまで達している。また、誰にも送金を止めさせないことを目的とした Bitcoin[17] の登場を皮切りに、ブロックチェーンを用いた様々な分散型台帳技術を用いたサービスが出現している。こうしたサービスは P2P と呼ばれる通信技術を用いており、世界中のコンピューターがそのノードとしてシステムの運用を担っていため、特定の国家がそのサービスを規制することは極めて困難である。グローバル化や新技術の登場は今後も止まることなく、インターネットは国家という枠組みを超越した社会インフラへ進化していくだろう。「地球規模 OS」[21] のような地球規模で資源を抽象化して共有可能にするプラットフォームが当たり前に存在する未来がやってくるかもしれない。

そうした未来に向かって必ず衝突するのは、国家を超越したインターネット上に存在するサービスの正当性をどのように保証するかという問題である。先に述べたとおり、特定の国家の法によってサービスの運営母体を規制しサービスの正当性を保証させることは困難になりつつある。また、Bitcoin ブロックチェーンにおいては、Proof of Work と呼ばれる技術によって確率的にその正当性を保証しようとしているが、そのためだけに世界中で大量の計算リソースが消費され続けているのが現実である。もちろんそれ以外に手段が存在し得ない可能性もあるものの、これが国家を超越した社会インフラの正当性を保証するためにベストな方法であるとは容易に納得し難い。

我々はこの問題を解決する糸口は社会契約の理論研究にあると考えている。かつて中世の社会契約論者が説明したように、国家の法が国民の社会契約によって成立するのであれば、国家を超越したインターネット上の法はインターネットユーザーによる社会契約によって成立するのではないだろうか。仮に国家を超越したインターネット上の法が成立するのならば、それによってサービスの正当性を保証すればよい。こうしたアイディアが本研究が社会契約を取り扱うモチベーションである。

1.2 本研究の貢献

本研究では、我々はインターネットユーザーの集団を想定して社会契約の理論を再構築する。当然のごとく、彼らはコンピューターを用いてインターネットでやり取りをするため、社会契約のモデルは (コンピューターで) 計算可能なレベルまで抽象化されている必要がある。そこに至る最大の障壁は強制執行力と呼ばれるプレイヤーの利得の分配を強制執行する力の存在である。これまでの社会契約の研究においては、この強制執行力が成員のどのような振る舞いによって生じるのか明確にできていないため、社会契約全体を計算可能なレベルまで抽象化することは困難であった。我々はこの強制執行力を自己組織化させる成員達の振る舞いを示すことで、社会契約全体を計算可能なレベルでの抽象化すること示す。また、マルチエージェントシミュレーションによってそのモデルを検証することで、成員の性質と社会契約の成立の関係性を明らかにする。

1.3 本論の構成

本論の構成は次の通りである。第 2 章では本論を読みすすめるにあたって必要となる前提知識と先行研究について述べる。第 3 章では本研究で取り扱う問題について詳細に定義する。第 4 章では先に定義した問いに対する我々の仮説とそれを示すために必要となる 3 つの補題について述べるとともに、それらの全体的な検証方法について述べる。続く 3 つの章では、それぞれの補題についての具体的な検証方法と検証結果について述べる。第 8 章では全ての検証を通してわかったことについてまとめるとともに、そこから生まれた新たな疑問と今後の研究アイディアについて述べる。

第2章　背景

2.1　社会契約とは

社会契約とは、ある集団においてその成員達が合意された約束を必ず履行する状態に至るプロセスである。中世において、社会契約はホッブズやロック、ルソーによって議論され、なぜ国家において国民が法に従うのかを論理的に説明しようと試みることで王権神授説を否定する主張として注目された。現代においては、ロールズやハーサニーが中世の社会契約論の一般化を試み、その社会契約のプロセスの果に導きだされる公正な正義とはどういったものかを議論した。それ以降、社会契約は倫理や正義論といった概念とともに、ゲーム理論を用いて分析されるようになり、機能主義、交渉理論、進化ゲーム理論の 3 つのアプローチがなされている。[27]

2.2　スタグハントゲーム

		$hunter_2$	
		鹿	野兎
$hunter_1$	鹿	(2,2)	(0,1)
	野兎	(1,0)	(1,1)

表 2.1: スタグハントゲームの利得表

スタグハントゲームとは、ルソーの「人類不平等起源論」[19] に登場する「鹿狩りの寓話」をモデリングした非協力戦略型ゲームである。[24] 二人のハンターが協力して鹿を狩るか、相手を裏切って野兎を狩るかを選択するが、鹿は 2 人で協力しなければ狩ることができず、1 人だけで狩ろうとすると何も得ることができない。このゲームは代表的な囚人のジレンマゲームであり、非協力解 (両者が野兎を狩る戦略をとる解) がナッシュ均衡になることが知られている。こうした囚人のジレンマゲームは社会契約が成立する原理を説明する鍵だと考えられている。[24]

2.3　機能主義的なアプローチ

Edna Ullmann-Margalit は、同様の囚人のジレンマゲームを持ち出し、道徳的な規範の機能は、各プレイヤーが合理的な戦略決定の果てに非協力解に陥るのを防ぐことであると説明した。[8] Mackie の研究においても、同様に道徳の機能は合理性の失敗を防ぐことにあるとされる。[16] しかしながら、こうしたアプローチは道徳的な規範の機能を説明する一方で、なぜプレイヤーがその規範に従うのかについては説明がなされていない。

2.4 交渉理論を用いたアプローチ

交渉理論を用いたアプローチとしては、Harsanyi 1955[13] や Rawls 1971[18]、Gauthier 1986[?] などが挙げられる。彼らは協力解が効率的な解であると考え、交渉理論を用いてどのような協力解が選択されるのかを説明しようした。しかしながら、このアプローチは選択可能な社会契約の集合からそれぞれの交渉解の要件に沿った解が選択されるのかを説明するものに過ぎず、具体的にどういったプロセスでその解に至るのかを説明しない。

2.5 進化ゲーム理論を用いたアプローチ

進化ゲーム理論を用いたアプローチは、道徳的な規範がどのようなプロセスで出現し維持されるのかを説明するものである。複数の研究で [26][3][4][23] 限定合理的な (2.6 節) エージェントの間でも道徳的な規範が出現することが示されており、こうした規範は繰り返されるエージェントの相互作用の中で創発されるものだとしている。それゆえ、安定した規範が必ずしもパレート効率だとは限らず、交渉理論を用いたアプローチが仮定しているような効率性と道徳性の関係は存在しないといえる。[27]

2.6 限定合理性

限定合理性とは、意思決定主体が認知能力の限界によって限定された合理性しか発揮することができない性質である。[22] ゲーム理論の文脈では、「自己利得を最大化する合理的なプレイヤー」という仮定の上では説明が困難だった実社会の協力的な行動を説明するために用いられている。[28] 最後通牒ゲームの実験はその最もたる例であり、実社会のプレイヤーは相手よりも少ない取り分を提示されると自身の利得が 0 になるにも関わらず報復的な戦略を選ぶことが知られている。[12]

2.7 強制執行力

これまでの社会契約の研究では、強制執行力と呼ばれる「合意された利得の分配を強制的に執り行っている力」の存在が暗黙的に仮定されている。先に紹介したスタグハントゲームでいうところの、二人で協力して鹿を狩った場合にその鹿を二人で分け合うという取り決めを守らせている力である。Rawls の「自然の義務」[18] や Harsanyi の「道徳的コミットメント」[13] などがこれに当たるとされる。[5]

2.8 外部の強制執行力

強制執行力のうち、集団の外部の機関によってもたらされる力を外部の強制執行力と呼ぶ。外部の強制執行力が存在する場合、そこには別の社会契約が存在していることになる。Binmore は交渉理論と進化ゲーム理論を用いてこの外部の強制執行力が存在する場合と存在しない場合の社会契約について分析し、下記のような結論を導き出した。[5]

外部の強制執行力が存在する場合、合理的なプレイヤーが合意する可能性のある契約はどれも無限回の繰り返しゲームの均衡結果となり得る。

外部の強制執行力が存在しない場合、過去の歴史が決定する社会指標を利用した平等主義的 (ロールズ的) 交渉解によって解決される。

2.9 複雑系

複雑系とは相互作用しうる複数の要素によって構成されるシステムである。複雑系全体の振る舞いは各構成要素の振る舞いによって決定論的に決まるが、その関係性が非線形的であり些細なパラメーターの変化で結果が大きく異なってしまうため、個々の振る舞いから全体の挙動を予測することは困難である。還元主義的なアプローチはシステム全体の振る舞いをシステムを分解することで理解しようとするのに対し、複雑系のアプローチは構成要素の振る舞いの変化とそれに伴うシステム全体の振る舞いの変化を観察することで理解しようとする。

本論では、社会契約を成立させようとする集団を、成員によって構成される複雑系として捉え、彼らの振る舞いの変化が社会契約の成立にどのように影響をもたらすのかを解析する。

2.10 自己組織化

自己組織化とは、システム全体を俯瞰できない構成要素の振る舞いによって、全体として秩序だった振る舞いがなされる現象である。2.5 節で述べたような道徳的な規範が生じる現象はこの自己組織化の一例だと考えられる。本論では、強制執行力を社会契約の複雑系の中で自己組織化された現象として設計する。設計にあたっては次節で紹介するビザンチン将軍問題とその解決策を用いる。

2.11 ビザンチン将軍問題

ビザンチン将軍問題とは、分散システムの構成し相互に通信しあうノード群において、それぞれのノードが本体の故障または故意によって偽の情報を伝達したり何も情報を伝達しない可能性がある場合に、全ての正常なノードが単一の値を共有することができるかを問う問題である。[15] 名称の通り、ビザンチン帝国の将軍たちが指令を共有する問題として記述されており、具体的には次のようなものである。

ビザンチン帝国の将軍たちが 1 つの都市を包囲しており、「攻撃」か「撤退」か合意したいと考えている。一部の将軍たちは「攻撃」を提案し、他は「撤退」を提案するかもしれないが、一部の将軍だけで攻撃すると失敗してしまう。将軍たちは、それぞれ離れた場所にいるため、メッセンジャーを相互に送って自分の指令を伝えようとするが、全ての将軍が誠実とは限らず、中には裏切り者もいて意見を分断させようとするかもしれない。ここで自身の指令 (「攻撃」か「撤退」) を伝えようとしている将軍を司令官、他の全ての将軍を副官としたとき、IC1 と IC2 を同時に達成する方法はあるだろうか。

IC1. すべての誠実な副官は同じ指令に従う。

IC2. 司令官が誠実な場合、全ての誠実な副官は彼の送った指令に従う。

2.12 署名されたメッセージによる解決策

この問題にはいくつかの解決策が存在しているが、ここでは問題が提起されたの論文の中で取り上げられている「署名されたメッセージによる解決策」を紹介する。それによれば、裏切り者の人数を m としたとき、A1〜A5 の 5 つの仮定の上で、$m+1$ 人以上の将軍がいれば下記のアルゴリズムで IC1 と IC2 を同時に満たせることが証明されている。将軍の人数を n としたとき、$i \in \{1, ..., n-1\}$ であり、$lineutenant_i$ は i 番目の将軍を指す。任意の値 w に対して、$w:i$ は i 番目の将軍によって署名がついた値である。$w:0$ の場合は司令官の将軍の署名がついた値を指す。

2.12.1 仮定

A1 送信されたすべてのメッセージは正しく到達する

A2 メッセージの受信者は誰が送信したのかわかる

A3 メッセージが届かないことを検知できる

A4 誠実な将軍の署名は偽造できず、署名されたメッセージの内容が変更されても、それを検知することができる。

A5 誰でも将軍の署名の信憑性を検証することができる。

2.12.2 関数 $choice(V)$

関数 $choice(V)$ は集合 V を引数にとって指令 (「攻撃」か「撤退」) を返す関数である。

1. もし集合に単一の指令 v しか存在しなければ、$choice(V) = v$ とする。
2. $choice(\emptyset) = RETREAT$ とする。$\emptyset$ は空集合。
3. もとの論文で抜けてるけど、V が 2 つあるときは $RETREAT$?

2.12.3 アルゴリズム $SM(m)$

1. $V_i = \emptyset$ として初期化する。($\emptyset$ は空集合)
2. 司令官は彼の値を全ての副官に署名して送る。
3. 各 i について、
   (a) もし $lineutenant_i$ が $v:0$ という形式のメッセージを受け取り、まだ何の命令も受けていない場合は、
      i. $lineutenant_i$ は V_i を v にする。
      ii. $lineutenant_i$ は他のすべての中尉にメッセージ $v:0:i$ を送ります。
   (b) もし $lieutenant_i$ が $v:0:j_1:...:j_k$ という形式のメッセージを受け取り、v が集合 V_i に入っていない場合は
      i. $lieutenant_i$ は v を V_i に追加する。
      ii. もし $k < m$ であれば、$lieutenant_i$ は $j_1, ..., j_k$ 以外のすべての副官に $v:0:j_1:...:j_k:i$ というメッセージを送る。
4. 各 i について。$lineutenant_i$ がこれ以上メッセージを受け取らない場合、$lieutenant_i$ は命令 $choice(V_i)$ に従う。

2.13 マルチエージェントシミュレーション

マルチエージェントシミュレーションとは、与えられた方策に従って振る舞う複数のエージェントとそのエージェント達を内包する環境を定義し、計算機によってそれらのエージェントの振る舞いをシミュレーションすることでエージェント達の振る舞いによる相互作用をシミュレートする手法である。本論では、分散システムとして設計されたシステムの検証のために用いる。

第3章　問題提起

3.1　本論の問題提起

これまでのゲーム理論を用いた社会契約の研究では、限定合理的な成員の繰り返しゲームによって道徳的な規範が創発され社会契約が成立することがわかっている [23][25]。また、社会契約が成立した際、どのような解が選択されるのかについても研究されてきた。[18][13][10] しかしながら、こうした議論の中では強制執行力の存在が仮定されており、社会契約全体をモデリングするにはこの力が集団の内部で生じるメカニズムを解明する必要がある。

これについて Binmore の 2005 年の研究 [5] では、「外部の強制執行力が存在しない場合、社会の過去の歴史が決定する社会指標を用いることで決定される平等主義的交渉解が選択される」という結論に至っているが、この研究は還元主義的なアプローチを用いているため社会契約に必要な歴史の定義や社会指標の計算方法、成員達の具体的な振る舞いを明確に記述できていない。また、この研究はすべての成員が各成員の戦略の選好を共有知識として持つという仮定の上に成り立っているが、こうした仮定は集団の人数が多い場合やインターネット上など、直接的に他の成員の行動を観察できない環境では現実的ではない。

そこで本研究では、集団の外部に強制執行力が存在しない場合に社会契約が成立しうるのかという問いに取り組み、その問を解く過程で社会契約の成立に必要となる過去の歴史と社会指標の決定方法、成員達の振る舞いを計算可能なレベルで明確にする。ここでの社会契約の成立とは、その集団の成員達が必ず合意された約束を履行する状態に至ることである。

3.2　問題解決の要件

先の問題が解決されるためには、成員達が合意された約束を遵守する状態に至ることを示す他に、下記の 3 つの要件が満たされている必要がある。

3.2.1　歴史を定義

第 1 に、社会契約の成立のために必要な歴史とはいかなるものかを定義する必要がある。これは社会指標を計算するために各成員が記録すべき過去のある集団の何らかの状態である。

3.2.2　社会指標の決定方法

第 2 に、歴史から社会指標を計算する方法を定義する必要がある。合意された約束が遵守されるようにするため、この社会指標は約束を履行する成員の社会指標が上がり、約束を保護にする成員の社会指標が下がるように設計する必要がある。

3.2.3 成員の振る舞い

第 3 に、各成員の振る舞いによってのみ、歴史が共有されて各成員の社会指標を計算でき、その振る舞いを記述可能である必要がある。これは外部の強制執行力が存在しない場合、あらゆる歴史も計算された社会指標も、その正当性を保証した状態で外部に記録共有することができないためである。

第4章　仮説と検証方法

4.1　本論の仮説

我々は、先の問について、外部の強制執行力が存在しないとき、一部の成員が誠実に振る舞っていれば社会契約は成立すると考えている。本論では、この仮説を示すために下記の３つ仮説を補題として扱い順に検証する。

補題１　外部に強制執行力が存在するとき、全ての成員が完全に合理的ならば社会契約は成立しない。

補題２　外部に強制執行力が存在するとき、一部の成員が限定合理的ならば社会契約は成立する。

補題３　成員の振る舞いによって補題２の強制執行力を自己組織化することができる。

4.2　検証方法

本論では、補題１～3 を順に検証することで、仮説を示す。その手順について概要をここで述べる。

まず、成員の行動を観察できない外部の強制執行力を「評判システム」と定義した上で議論を進め、補題１を示す。「評判システム」とは報告された約束の結果に基づいて各成員の評判スコアを決定するシステムである。また、そのシステムを用いた任意の約束を結ぶ２人の「約束・評判ゲーム」について考える。これは一方 ($promisor$) が約諾し、もう一方 ($reporter$) がその約束の結果 (「成功」か「失敗」) を「評判システム」に報告する非協力戦略型ゲームである。このゲームに参加するプレイヤーが完全に合理的な場合、両者がとりうる各戦略の利得を比較することで、彼らの約束が真に成功する ($promisor$ が約束を履行し、$reporter$ が「成功」を報告する) 条件を導く。全ての成員が完全に合理的な場合、「評判システム」が報告された約束の結果から、その条件を満たす評判スコアを決定することができないことを示す。

次に、補題１を踏まえて、全ての成員が約束を反故にされた場合に「失敗」を報告する「告発する約束・評判ゲーム」について考えることで補題２を示す。そして、補題１と同様に、成員がとりうる各戦略の利得を比較することで、この「告発する約束・評判ゲーム」において、彼らの約束が真に成功する条件を導く。「告発する約束・評判ゲーム」においては、「評判システム」が報告された約束の結果 (「成功」か「失敗」) から、その条件を満たす評判スコアを決定することができることを示す。その条件を満たす「評判システム」の詳細を定義し、戦略の限定されない通常の「約束・評判ゲーム」において、定義した評判システムが機能するかマルチエージェントシミュレーションを用いて検証する。その結果から、集団を構成する成員達の性質によっては、全ての約束が成功する状態に至ることを示す。

最後に、各成員が「評判システム」を所有している場合について考えることで、補題３を示す。ビザンチン将軍問題の署名付きの解決策を用いて、「評判システム」を各成員の振る舞いによる分散システムとして設計可能であることを示す。補題２の条件に基づいて実装された「評判システム」を同様に各成員の振る舞いとして定義し、マルチエージェントシミュレーションを用いて検証する。その結果から、外部の強制執行力が存在しない場合でも、成員の振る舞いによって強制執行力を自己組織化させることが可能であり、集団を構成する成員達の性質によっては、全ての約束が成功する状態に至ることがわかる。

また、補題３の実験結果から、「評判システム」が成員達の振る舞いによって自己組織化された分散システムとして機能し、集団を構成する成員達の性質によっては、「約束-評判ゲーム」で全ての約束が成功する状態に至ることが示される。これにより、先の仮説が立証され、社会契約の成立に必要となる歴史の定義や社会指標の決定方法、具体的な成員の振る舞い、その振る舞いに従う成員の人数と社会契約の成立の関係性を明らかにすることができる。

第5章　補題1の検証

本章では、4.1 節で定義した補題 1 を検証する。

5.1　補題 1

外部に強制執行力が存在するとき、全ての成員が完全に合理的ならば社会契約は成立しない。

5.2　検証方法

成員の行動を観察できない外部の強制執行力として「評判システム」(5.3 節) の存在を仮定し、2 人の成員が、そのシステムを用いて約束を交わす「約束・評判ゲーム」(5.4) について考える。このとき、各成員が約束によって生じる価値と「評判スコア」の合計を最大化しようとする場合、約束が履行されるような「評判スコア」を「評判システム」から決定できないことを示す。

5.3　評判システム

「評判システム」とは、初期の各成員の「評判スコア」と報告された約束の記録に基づいて、各成員の「評判スコア」を決定するシステムである。約束の記録とは、約諾者が約束を履行したかについての情報であり、約諾者、報告者、結果 (「成功」か「失敗」) からなる。このシステムから成員の行動を観察することはできないため、約諾者が真に約束を履行したか否かと報告された結果が同じとは限らない。このシステムは、「約束・評判ゲーム」(5.4 節) において、強制執行力としての役割を果たす。

5.4　約束・評判ゲーム

「約束・評判ゲーム」とは、約諾者 (*promisor*) と報告者 (*reporter*) の 2 人によって行われるゲームである。約諾者は 2 者間で合意された約束を履行する、もしくは反故にする。それに対して報告者は約束の結果 (「成功」か「失敗」) を決定し、約束の記録を「評判システム」に報告する。

step1 *promisor* は合意された約束を履行する、もしくは反故にする.

step2 *reporter* は「成功」か「失敗」を「評判システム」に報告する.

5.4.1 展開型ゲーム

これは図 5.1 のゲームの木のような展開型ゲームとして表せる．各変数については下記の通りである。

step1 で *promisor* が約束を履行するか反故にするかと、**step2** で *reporter* が「成功」を報告するか「失敗」を報告するかで 4 つの結果がある。また，「評判システム」からは *promisor* と *reporter* の行動を観察できないため，**step2** での *reporter* の報告に基づいて *promisor* と *reporter* の「評判スコア」が決定しなければならない．それ故、①と③，②と④はそれぞれ $c_k(k$ は任意) を除いて同じ利得でなくてはならない．

各変数の定義

c_{p1} … 約束が履行された場合の *promisor* の「評判スコア」の変化量以外の効用

c_{p2} … 約束が反故にされた場合の *promisor* の「評判スコア」の変化量以外の効用

c_{r1} … 約束が履行された場合の *reporter* の「評判スコア」の変化量以外の効用

c_{r2} … 約束が反故にされた場合の *reporter* の「評判スコア」の変化量以外の効用

r_{ps} … 「成功」が報告された場合の *promisor* の「評判スコア」の変化量

r_{pf} … 「失敗」が報告された場合の *promisor* の「評判スコア」の変化量

r_{rs} … 「成功」が報告された場合の *reporter* の「評判スコア」の変化量

r_{rf} … 「失敗」が報告された場合の *reporter* の「評判スコア」の変化量

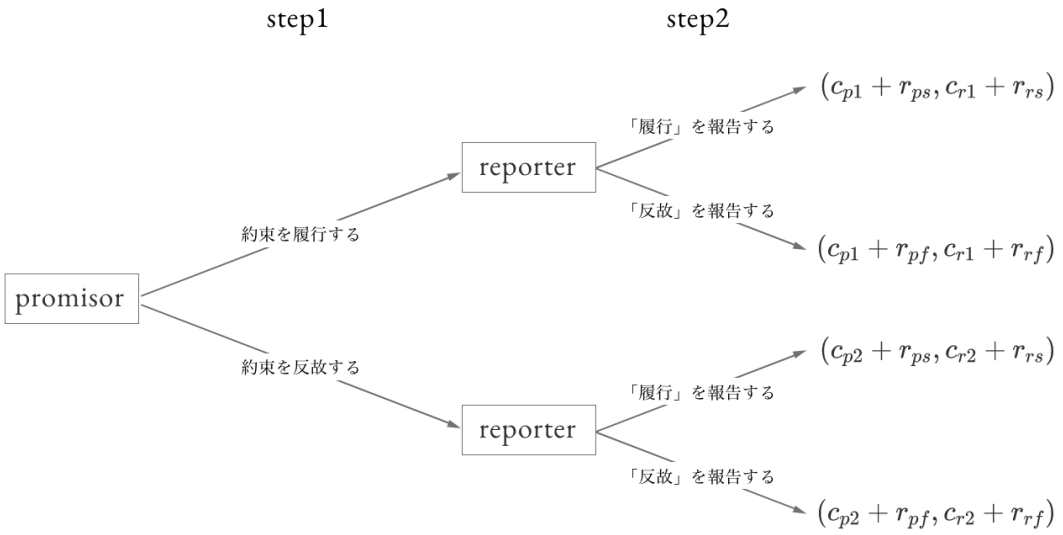

図 5.1: 「約束・評判ゲーム」のゲーム木

5.4.2 各プレイヤーの戦略

この展開型ゲームにおける *promisor* と *reporter* の行動は、下記のような戦略として表せる。

promisor の戦略

s_{p1} … 約束を履行する

s_{p2} … 約束を反故にする

reporter の戦略

s_{r1} … *promisor* が約束を履行した場合は「成功」，反故にした場合は「失敗」を報告する

s_{r2} … *promisor* が約束を履行した場合は「成功」，反故にした場合は「成功」を報告する

s_{r3} … *promisor* が約束を履行した場合は「失敗」，反故にした場合は「失敗」を報告する

s_{r4} … *promisor* が約束を履行した場合は「失敗」，反故にした場合は「成功」を報告する

5.4.3 非協力戦略型ゲーム

これらの戦略を用いて、先に述べた展開型ゲームは表 5.1 のような非協力戦略型ゲームとして書き換えられる．

		Reporter			
		s_{r1}	s_{r2}	s_{r3}	s_{r4}
Promisor	s_{p1}	$(c_{p1}+r_{ps},$ $c_{r1}+r_{rs})$	$(c_{p1}+r_{ps},$ $c_{r1}+r_{rs})$	$(c_{p1}+r_{pf},$ $c_{r1}+r_{rf})$	$(c_{p1}+r_{pf},$ $c_{r1}+r_{rf})$
	s_{p2}	$(c_{p2}+r_{pf},$ $c_{r2}+r_{rf})$	$(c_{p2}+r_{ps},$ $c_{r2}+r_{rs})$	$(c_{p2}+r_{pf},$ $c_{r2}+r_{rf})$	$(c_{p2}+r_{ps},$ $c_{r2}+r_{rs})$

表 5.1: 「約束・評判ゲーム」の利得票表

5.5 不正が防止される条件

表 5.1 より，「約束・評判ゲーム」において、約束が履行されるためには、*promisor* と *reporter* のとる戦略組が (s_{p1}, s_{r1}) もしくは (s_{p1}, s_{r2}) のいずれかになる必要がある．各プレイヤーが戦略 s をとったときの利得を R とし、その期待値を $E(R|s)$ とする。全てのプレイヤーが完全に合理的な場合、(s_{p1}, s_{r1}) か (s_{p1}, s_{r2}) のいづれかの戦略組に帰着させるためには，

条件① $E(R|s_{p1}) > E(R|s_{p2})$ かつ $E(R|s_{r1}) > \max\{E(R|s_{r2}), E(R|s_{r3}), E(R|s_{r4})\}$

条件② $E(R|s_{p1}) > E(R|s_{p2})$ かつ $E(R|s_{r2}) > \max\{E(R|s_{r1}), E(R|s_{r3}), E(R|s_{r4})\}$

のいづれかを満たす $(r_{ps}, r_{pf}, r_{rs}, r_{rf})$ の組を「評判システム」から決定できる必要がある．本章ではこれが不可能であることを示す．

5.6 命題

条件①か条件②のいづれかを満たす $(r_{ps}, r_{pf}, r_{rs}, r_{rf})$ の組を「評判システム」から決定することはできない．

5.7 証明

各プレイヤーが戦略 s_k をとる確率を p_k とする.

$$
\begin{aligned}
0 \leq p_k \leq 1 \\
p_{p1} + p_{p2} = 1 \\
p_{r1} + p_{r2} + p_{r3} + p_{r4} = 1
\end{aligned} \tag{5.1}
$$

各戦略の期待利得

promisor と *reporter* の各戦略の期待利得は次のように表せる.

$$
\begin{aligned}
E(R|s_{p1}) &= p_{r1}(c_{p1} + r_{ps}) + p_{r2}(c_{p1} + r_{ps}) + p_{r3}(c_{p1} + r_{pf}) + p_{r4}(c_{p1} + r_{pf}) \\
&= c_{p1} + p_{r1}r_{ps} + p_{r2}r_{ps} + p_{r3}r_{pf} + p_{r4}r_{pf} \because (5.1) && (5.2) \\
E(R|s_{p2}) &= p_{r1}(c_{p2} + r_{pf}) + p_{r2}(c_{p2} + r_{ps}) + p_{r3}(c_{p2} + r_{pf}) + p_{r4}(c_{p2} + r_{ps}) \\
&= c_{p2} + p_{r1}r_{pf} + p_{r2}r_{ps} + p_{r3}r_{pf} + p_{r4}r_{ps} \because (5.1) && (5.3) \\
E(R|s_{r1}) &= p_{p1}(c_{r1} + r_{rs}) + p_{p2}(c_{r2} + r_{rf}) && (5.4) \\
E(R|s_{r2}) &= p_{p1}(c_{r1} + r_{rs}) + p_{p2}(c_{r2} + r_{rs}) && (5.5) \\
E(R|s_{r3}) &= p_{p1}(c_{r1} + r_{rf}) + p_{p2}(c_{r2} + r_{rf}) && (5.6) \\
E(R|s_{r4}) &= p_{p1}(c_{r1} + r_{rf}) + p_{p2}(c_{r2} + r_{rs})
\end{aligned}
$$

条件①が成り立たないことの証明

条件①が成り立たないことを示すために、その必要条件である下記の 2 つの条件について考える。

$$
E(R|s_{r1}) > E(R|s_{r2}) \tag{5.7}
$$

$$
E(R|s_{r1}) > E(R|s_{r3}) \tag{5.8}
$$

(5.7) を満たすためには、

$$
\begin{aligned}
& E(R|s_{r1}) > E(R|s_{r2}) \\
\therefore\ & p_{p1}(c_{r1} + r_{rs}) + p_{p2}(c_{r2} + r_{rf}) > p_{p1}(c_{r1} + r_{rs}) + p_{p2}(c_{r2} + r_{rs}) \because (5.4)(5.5) \\
\therefore\ & p_{p2}r_{rf} - p_{p2}r_{rs} > 0 \\
\therefore\ & p_{p2}(r_{rf} - r_{rs}) > 0
\end{aligned}
$$

つまり、

$$
p_{p2} > 0
$$

かつ

$$
0 > r_{rs} - r_{rf} \tag{5.9}
$$

を満たす必要がある。

(5.8) を満たすためには、

$$
\begin{aligned}
& E(R|s_{r1}) > E(R|s_{r3}) \\
\therefore\quad & p_{p1}(c_{r1} + r_{rs}) + p_{p2}(c_{r2} + r_{rf}) > c_{p2} + p_{r1}r_{pf} + p_{r2}r_{ps} + p_{r3}r_{pf} + p_{r4}r_{ps}) \because (5.4)(5.6) \\
\therefore\quad & p_{p1}r_{rs} + p_{p2}r_{rf} > p_{p1}r_{rf} + p_{p2}r_{rf} \\
\therefore\quad & p_{p1}(r_{rs} - r_{rf}) > 0
\end{aligned}
$$

つまり，

$$p_{p1} > 0$$

かつ

$$r_{rs} - r_{rf} > 0 \tag{5.10}$$

を満たす必要がある。

ここで (5.9) と (5.10) を同時に満たすことはできないため，条件①は成り立たない.

条件②が成り立たないことの証明

次に、条件②の必要条件である

$$E(R|s_{p1}) > E(R|s_{p2}) \tag{5.11}$$

について考える。

(5.11) を満たすためには，

$$
\begin{aligned}
& E(R|s_{p1}) > E(R|s_{p2}) \\
\therefore\quad & c_{p1} + p_{r1}r_{ps} + p_{r2}r_{ps} + p_{r3}r_{pf} + p_{r4}r_{pf} > c_{p2} + p_{r1}r_{pf} + p_{r2}r_{ps} + p_{r3}r_{pf} + p_{r4}r_{ps} \\
\therefore\quad & c_{p1} + p_{r1}r_{ps} + p_{r4}r_{pf} > c_{p2} + p_{r1}r_{pf} + p_{r4}r_{ps} \\
\therefore\quad & c_{p1} - c_{p2} + p_{r1}r_{ps} + p_{r4}r_{pf} + p_{r1}r_{pf} - p_{r4}r_{ps} > 0 \\
\therefore\quad & p_{r1}(r_{ps} - r_{pf}) - p_{r4}(r_{ps} - r_{pf}) + c_{p1} - c_{p2} > 0 \\
\therefore\quad & (p_{r1} - p_{r4})(r_{ps} - r_{pf}) + c_{p1} - c_{p2} > 0 \\
\therefore\quad & (p_{r1} - p_{r4})(r_{ps} - r_{pf} + \frac{c_{p1} - c_{p2}}{p_{r1} - p_{r4}}) > 0
\end{aligned}
$$

を満たす必要がある．つまり，

$p_{r1} > p_{r4}$ のとき，

$$r_{ps} - r_{pf} + \frac{c_{p1} - c_{p2}}{p_{r1} - p_{r4}} > 0$$
$$\therefore \quad r_{ps} - r_{pf} > \frac{c_{p2} - c_{p1}}{p_{r1} - p_{r4}}$$

$p_{r1} < p_{r4}$ のとき，

$$r_{ps} - r_{pf} + \frac{c_{p1} - c_{p2}}{p_{r1} - p_{r4}} < 0$$
$$\therefore \quad r_{ps} - r_{pf} < \frac{c_{p2} - c_{p1}}{p_{r1} - p_{r4}}$$

を満たせばよい.

ここで、$E(R|s_{r2}) > \max\{E(R|s_{r1}), E(R|s_{r3}), E(R|s_{r4})\}$ が成り立つと仮定する。

このとき全ての合理的なプレイヤーは戦略 s_r2 をとり、必ず「成功」を報告するため、「評判システム」から p_{r1} と p_{r4} を推定することはできない。

ゆえに、「評判システム」から条件②を満たすような (r_{ps}, r_{pf}) の組を決定できない.

以上より，条件①と条件②のいづれかを満たす $(r_{ps}, r_{pf}, r_{rs}, r_{rf})$ の組を「評判システム」から決定することはできない.

Q.E.D.

メモ：後で $c_{p1} - c_{p2} \neq 0$ の仮定を追加？

第6章　補題2の検証

本章では、4.1 節で定義した補題 2 を検証する。

6.1　補題2

外部に強制執行力が存在するとき、一部の成員が限定合理的ならば社会契約は成立する。

6.2　提案手法

不正行為にあった場合に必ず「失敗」を報告する行動を「告発」と呼び、「告発」するプレイヤーのみが参加する「約束・評判ゲーム」を「告発する約束・評判ゲーム」とする。この「告発する約束・評判ゲーム」において、不正が防止される条件を導き出す。そして、その条件を満たす「評判システム」を通常の「約束・評判ゲーム」に適用し、様々な戦略をとるエージェントをランダムに用意したマルチエージェントシミュレーションを行ことで、「告発」するプレイヤーが一定以上存在していれば成員達に合意された約束を履行させることができることを示す。

6.3　告発

5.7 節の検証では、「評判システム」は *reporter* が戦略 s_{r1} と s_{r4} をとる確率 p_{r1} と p_{r4} を推定できないため、不正を防止する「評判スコア」の変化量の組 $(r_{ps}, r_{pf}, r_{ps}, r_{pf})$ を決定できなかった。これは *reporter* が報告する約束の結果と真の約束の結果が一致している保証がないためである。そこで、本論では、*promiser* が約束を反故にした場合に「失敗」を報告する行動を「告発」とする。この行動をとる成員が多数いれば、先の確率を近似することが可能であると考えられる。

6.4　告発する約束・評判ゲーム

「告発」する成員のみで行われる「約束・評判ゲーム」を「告発する約束・評判ゲーム」とする。このゲームにおいては、step1 で *promisor* が約束を履行しなかった場合、*reporter* は必ず「失敗」を報告するため、図 6.1 のゲームの木のような展開型ゲームとして表せる。また、5.4.2 節で示した *reporter* の戦略のうち s_{r2} と s_{r4} がとられないため、非協力戦略型ゲームとして表したときの利得は表 6.1 のようになる。注意すべき点としては、報告された約束の結果が次回以降のゲームで意味のある情報となるため、「成功」と「失敗」が報告された場合の将来の期待利得 ϵ と λ について考慮する必要がある。

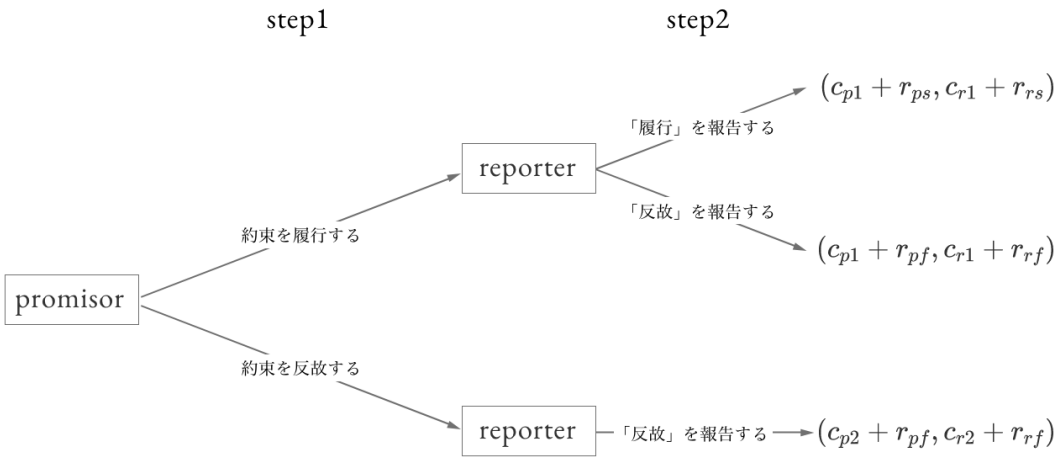

図 6.1: 「告発する約束・評判ゲーム」のゲーム木

		reporter	
		s_{r1}	s_{r3}
promisor	s_{p1}	$(c_{p1} + r_{ps} + \epsilon_p,$ $c_{r1} + r_{rs} + \epsilon_r)$	$(c_{p1} + r_{pf} + \lambda_p,$ $c_{r1} + r_{rf} + \lambda_r)$
	s_{p2}	$(c_{p2} + r_{pf} + \lambda_p,$ $c_{r2} + r_{rf} + \lambda_r)$	$(c_{p2} + r_{pf} + \lambda_p,$ $c_{r2} + r_{rf} + \lambda_r)$

表 6.1: 「告発する約束・評判ゲーム」の利得表

6.4.1 各プレイヤーの戦略

この展開型ゲームにおける *promisor* と *reporter* の行動は、下記のような戦略として表せる。

promisor の戦略

s_{p1} … 約束を履行する

s_{p2} … 約束を反故にする

reporter の戦略

s_{r1} … *promisor* が約束を履行した場合は「成功」，反故にした場合は「失敗」を報告する

s_{r3} … *promisor* が約束を履行した場合は「失敗」，反故にした場合は「失敗」を報告する

6.4.2 各変数の定義

ϵ_p … 「成功」が報告された場合の $promisor$ の将来期待利得

ϵ_r … 「成功」が報告された場合の $reporter$ の将来期待利得

λ_p … 「失敗」が報告された場合の $promisor$ の将来期待利得

λ_r … 「失敗」が報告された場合の $reporter$ の将来期待利得

c_{p2} … 約束が反故にされた場合の $promisor$ の「評判スコア」の変化量以外の効用

c_{p1} … 約束が履行された場合の $promisor$ の「評判スコア」の変化量以外の効用

c_{p2} … 約束が反故にされた場合の $promisor$ の「評判スコア」の変化量以外の効用

c_{r1} … 約束が履行された場合の $reporter$ の「評判スコア」の変化量以外の効用

c_{r2} … 約束が反故にされた場合の $reporter$ の「評判スコア」の変化量以外の効用

r_{ps} … 「成功」が報告された場合の $promisor$ の「評判スコア」の変化量

r_{pf} … 「失敗」が報告された場合の $promisor$ の「評判スコア」の変化量

r_{rs} … 「成功」が報告された場合の $reporter$ の「評判スコア」の変化量

r_{rf} … 「失敗」が報告された場合の $reporter$ の「評判スコア」の変化量

6.5 不正が防止される条件

表 5.1 より,「告発する約束・評判ゲーム」において、約束が履行され「成功」が報告されるためには、$promisor$ と $reporter$ のとる戦略組が (s_{p1}, s_{r1}) になる必要がある．各プレイヤーが戦略 s をとったときの利得を R とし、その期待値を $E(R|s)$ とする。全てのプレイヤーが完全に合理的な場合、戦略組 (s_{p1}, s_{r1}) に帰着させるためには,

条件③ $E(r|s_{p1}) > E(r|s_{p2})$ かつ $E(r|s_{r1}) > E(r|s_{r3})$

を満たす「評判スコア」の変化量の組 $(r_{ps}, r_{pf}, r_{rs}, r_{rf})$ を「評判システム」から決定できる必要がある．

6.5.1 各戦略のとられる確率

各プレイヤーが戦略 s_k をとる確率を p_k とする．

$$0 \leq p_k \leq 1$$

$$p_{p1} + p_{p2} = 1 \tag{6.1}$$

$$p_{r1} + p_{r3} = 1 \tag{6.2}$$

6.5.2 各戦略の期待利得

promisor と *reporter* の各戦略の期待利得は次のように表せる.

$$
\begin{aligned}
E(R|s_{p1}) &= p_{r1}(c_{p1} + r_{ps} + \epsilon_p) + p_{r3}(c_{p1} + r_{pf} + \lambda_p) \\
&= c_{p1} + p_{r1}(r_{ps} + \epsilon_p) + p_{r3}(r_{pf} + \lambda_p) \because (6.2) && (6.3) \\
E(R|s_{p2}) &= p_{r1}(c_{p2} + r_{pf} + \lambda_p) + p_{r3}(c_{p2} + r_{pf} + \lambda_p) \\
&= c_{p2} + r_{pf} + \lambda_p \because (6.2) && (6.4) \\
E(R|s_{r1}) &= p_{p1}(c_{r1} + r_{rs} + \epsilon_r) + p_{p2}(c_{r2} + r_{rf} + \lambda_r) && (6.5) \\
E(R|s_{r3}) &= p_{p1}(c_{r1} + r_{rf} + \lambda_r) + p_{p2}(c_{r2} + r_{rf} + \lambda_r) \\
&= p_{p1}c_{r1} + p_{p2}c_{r2} + r_{rf} + \lambda_r \because (6.1) && (6.6)
\end{aligned}
$$

6.5.3 *promiser* が s_{p1} をとる条件

$E(R|s_{p1}) > E(R|s_{p2})$ を満たすためには、

$$
\begin{aligned}
& E(R|s_{p1}) > E(R|s_{p2}) \\
\therefore\ & c_{p1} + p_{r1}(r_{ps} + \epsilon_p) + p_{r3}(r_{pf} + \lambda_p) > c_{p2} + r_{pf} + \lambda_p \\
\therefore\ & p_{r1}(r_{ps} + \epsilon_p) + (p_{r3} - 1)(r_{pf} + \lambda_p) > c_{p2} - c_{p1} \\
\therefore\ & p_{r1}(r_{ps} + \epsilon_p) - (1 - p_{r3})(r_{pf} + \lambda_p) > c_{p2} - c_{p1} \\
\therefore\ & p_{r1}(r_{ps} + \epsilon_p) - p_{r1}(r_{pf} + \lambda_p) > c_{p2} - c_{p1} \because (6.2) \\
\therefore\ & p_{r1}(r_{ps} - r_{pf} + \epsilon_p - \lambda_p) > c_{p2} - c_{p1} && (6.7)
\end{aligned}
$$

を満たせばよい.

6.5.4 *reporter* が s_{r1} をとる条件

$E(R|s_{r1}) > E(R|s_{r3})$ を満たすためには、

$$
\begin{aligned}
& E(R|s_{r1}) > E(R|s_{r3}) \\
\therefore\ & p_{p1}(c_{r1} + r_{rs} + \epsilon_r) + p_{p2}(c_{r2} + r_{rf} + \lambda_r) > p_{p1}c_{r1} + p_{p2}c_{r2} + r_{rf} + \lambda_r \\
\therefore\ & p_{p1}(r_{rs} + \epsilon_r) + p_{p2}(r_{rf} + \lambda_r) > r_{rf} + \lambda_r \\
\therefore\ & p_{p1}(r_{rs} + \epsilon_r) + (p_{p2} - 1)(r_{rf} + \lambda_r) > 0 \\
\therefore\ & p_{p1}(r_{rs} + \epsilon_r) - (1 - p_{p2})(r_{rf} + \lambda_r) > 0 \\
\therefore\ & p_{p1}(r_{rs} + \epsilon_r) - p_{p1}(r_{rf} + \lambda_r) > 0 \because (6.1) \\
\therefore\ & p_{p1}(r_{rs} - r_{rf} + \epsilon_r - \lambda_r) > 0 && (6.8)
\end{aligned}
$$

を満たせばよい。

6.5.5 戦略組 (s_{p1}, s_{r1}) に帰結する条件

$p_{r1} > 0$ かつ $p_{p1} > 0$ かつ $\epsilon_p > \lambda_p$ かつ $\epsilon_r > \lambda_r$ を仮定すると

$$r_{ps} - r_{pf} \geq \frac{c_{p2} - c_{p1}}{p_{r1}} \tag{6.9}$$

かつ

$$r_{rs} - r_{rf} \geq 0 \tag{6.10}$$

を満たせば,「告発する約束約束・評判ゲーム」で不正を防止することができる.

6.5.6 戦略 s_{p1} をとった割合と成功が報告される割合の関係

全ての成員は「告発」するため、任意の成員 i と j のこれまでの「約束・評判ゲーム」で s_{p1} をとってきた割合 $FulfillStrategyRate(i,j)$ と、「評判システム」に報告された約束の記録のうち $promisor$ が成員 i で $reporter$ が成員 j である記録の「成功」の割合 $ReportedSuccessRate(i,j)$ について、次の関係がいえる。

$$FulfillStrategyRate(i,j) \geq ReportedSuccessRate(i,j) \tag{6.11}$$

6.5.7 信頼度 P_i

ここで、成員 i が $promisor$ のときに s_{p1} をとる確率 p_{p1} を成員 i の信頼度 P_i と定義し, P_i を $FulfillStrategyRate(i,j)$ と $\sum_k^{\{1,2,...,n\}} w_k = 1$ を満たす任意の重み w_k ($k \in \{1,2,...,n\}$) を用いた荷重総和として次のように表す.

$$P_i \equiv p_{p1} \tag{6.12}$$

$$\equiv \sum_j^{\{1,2,...,n\}} w_j \cdot FulfillStrategyRate(i,j) \tag{6.13}$$

6.5.8 最低信頼度 T_i

また、$ReportedSuccessRate(i,j)$ に、同じ重み w_k を用いた荷重総和を最低信頼度 T_i と定義する.

$$T_i \equiv \sum_j^{\{1,2,...,n\}} w_j \cdot ReportedSuccessRate(i,j) \tag{6.14}$$

6.5.9 最低信頼度 T_i を用いた条件

(6.11) より、$P_i \geq T_i$ がいえるため、(6.9) について

$$\frac{c_{p2} - c_{p1}}{T_i} \geq \frac{c_{p2} - c_{p1}}{P_i} = \frac{c_{p2} - c_{p1}}{p_{r1}}$$

がいえる．ゆえに、

$$r_{ps} - r_{pf} \geq \frac{c_{p2} - c_{p1}}{T_i} \tag{6.15}$$

$$r_{rs} - r_{rf} \geq 0 \tag{6.16}$$

を満たせばよい、「告発する信用・評判ゲーム」で不正を防止できる。

ここで、「評判システム」から最低信頼度 T_i は既知のため、$p_{r1} > 0$ かつ $p_{p1} > 0$ かつ $\epsilon_p > \lambda_p$ かつ $\epsilon_r > \lambda_r$ を仮定した上で，$(r_{ps}, r_{pf}, r_{rs}, r_{rf})$ の組を決定できる。

6.6 評判システムの詳細

本節ではシミュレーションに用いる「評判システム」の仕様の詳細を紹介する。完全な実装については、GitHub のソースコードを参照。

6.6.1 約束の価値 C

$$C \equiv 1 \tag{6.17}$$

$$= c_{p2} - c_{p1} \tag{6.18}$$

$$= c_{r1} - c_{r2} \tag{6.19}$$

6.6.2 ReputationWeight

最低信頼度 T_i を求めるため、$ReportedSuccessRate(i, j)$ の荷重総和に用いる重み w_k を定義する．この重み w_k は、任意の成員 i が戦略 s_{p1} をとる確率 p_{p1} を推定する際に、成員 i と他の成員 j における $ReportedSuccessRate(i, j)$ をどの程度信頼するかを表している。本実験では、成員 j の「評判スコア」が全体に占める割合を ReputationWeightw_j と定義する。任意の成員 k の「評判スコア」を b_k としたとき、w_k は次のように表せる。(n は成員の人数)

$$w_k \equiv \frac{b_k}{\sum_k^{\{1,2,\ldots,n\}} b_k}$$

6.6.3 「成功」が報告された場合の「評判スコア」の変化

reporter が「成功」を報告した場合、*reporter* は約束の価値 C だけ減り、*promisor* は C だけ増加するものとする。

$$r_{ps} = C \tag{6.20}$$
$$r_{rs} = -C \tag{6.21}$$
$$r_{ps} + r_{rs} = 0 \tag{6.22}$$

6.6.4 EscrowCost

「失敗」が報告されたとき、*promisor* と *reporter* が失う「評判スコア」の合計を $EscrowCost$ とする。約束の価値 C にエスクロー係数 E を掛けたものとする.

$$\begin{aligned} EscrowCost &\equiv (r_{rs} - r_{rf}) + (r_{ps} - r_{pf}) && (6.23)\\ &= E \cdot C && (6.24) \end{aligned}$$

6.6.5 EscrowCost の負担比率

成員 i を *promiser*、j を *reporter* とする。「約束・評判ゲーム」を行うとき、$EscrowCost$ の負担比率を両者の最低信頼度 T_i を用いる.

$$(r_{ps} - r_{pf}) : (r_{rs} - r_{rf}) = T_j : T_i \tag{6.25}$$

6.6.6 EscrowCost の分配

「失敗」が報告されたときに $EscrowCost$ が消失すると、「評判スコア」の総量が下がり、約束の価値 C 総量を一定に保つために、*promisor* と *reporter* 以外の全てのプレイヤーに、$ReputationWeight$ に応じて $EscrowCost$ を分配する。*promisor* と *reporter* を含まないのは、分配によるインセンティブ設計への影響をなくすためである。

6.6.7 不正が防止される条件を満たす定義

成員 j を *promiser*、k を *reporter* とする。このとき、(6.15) と (6.16) を満たす $r_{ps} - r_{pf}$ と $r_{rs} - r_{rf}$ を下記のように定義する。

$$\begin{aligned} r_{ps} - r_{pf} &\equiv \frac{C}{T_i} && (6.26)\\ &\geq \frac{c_{p2} - c_{p1}}{T_i} \\ r_{rs} - r_{rf} &\equiv \frac{C}{T_j} && (6.27)\\ &\geq 0 \end{aligned}$$

6.6.8 残高の変化量の組 $(r_{ps}, r_{pf}, r_{rs}, r_{rf})$

上記の定義からエスクロー係数 E と残高の変化量の組 $(r_{ps}, r_{pf}, r_{rs}, r_{rf})$ は次のように求まる.
成員 i を *promisor*、j を *reporter* とする。

$$E = -\frac{T_i + T_j}{T_i T_j} \tag{6.28}$$

$$r_{ps} = 1 \tag{6.29}$$

$$r_{pf} = 1 - \frac{1}{T_j} \tag{6.30}$$

$$r_{rs} = -1 \tag{6.31}$$

$$r_{rf} = -1 - \frac{1}{T_i} \tag{6.32}$$

6.7 実験方法

「評判システムの詳細」に基づく「評判システム」と次の 8 タイプのエージェントから重複問わずランダムに選んだ 8 体のエージェントを用意し試行を実施する。これをタイプ A のエージェントが 0〜7 体を占める場合について、それぞれ 8000 回づつ繰り返し、エージェントの構成と step13 で求まる「報告された成功率」と「真の成功率」を記録する。

6.7.1 エージェントの種類

下記の A H の 8 タイプのエージェントを用意する。

A 約束を履行し、*promisor* が約束を履行したとき「成功」、反故にしたとき「失敗」を報告する。

B 約束を履行し、*promisor* が約束を履行したとき「成功」、反故にしたとき「成功」を報告する。

C 約束を履行し、*promisor* が約束を履行したとき「失敗」、反故にしたとき「成功」を報告する。

D 約束を履行し、*promisor* が約束を履行したとき「失敗」、反故にしたとき「失敗」を報告する。

E 約束を反故にし、*promisor* が約束を履行したとき「成功」、反故にしたとき「失敗」を報告する。

F 約束を反故にし、*promisor* が約束を履行したとき「成功」、反故にしたとき「成功」を報告する。

G 約束を反故にし、*promisor* が約束を履行したとき「失敗」、反故にしたとき「成功」を報告する。

H 約束を反故にし、*promisor* が約束を履行したとき「失敗」、反故にしたとき「失敗」を報告する。

6.7.2 試行

step 1 時刻 t を 0 とする。

step 2 「評判システム」の各エージェントの初期の「評判スコア」を 8 とする。

step 3 全てのエージェントが *promisor* と *reporter* として 1 度ずつ総当りする順序を決定する。(順序の長さは 56 となる)

step 4 時刻 t を 1 進める。

step 5 step 3 で決定した順序を周期として、*promisor* と *reporter* を決定する。

step 6 「評判システム」は「成功」と「失敗」が報告された場合の *promisor* と *reporter* の「評判スコア」を計算する。

step 7 *promisor* は自身の戦略に基づいて約束を履行するか反故にする。

step 8 *reporter* は自身の戦略と step 6 の *promisor* の行動に基づいて結果を決定する。

step 9 *reporter* は決定した結果を「評判システム」に報告する。

step 10 「評判システム」は step6 で計算した「評判スコア」がいずれの場合にも 0 未満にならない場合、reporter から報告された結果を記録する。

step 11 step6 で計算した「評判スコア」がいずれの場合にも 0 未満にならない場合、真の結果を記録する。

step 12 時刻 t が 1120 未満なら、step4 に戻る。

step 13 過去 56 回の約束・評判ゲームにおいて、step10 と 11 で記録された結果を集計し、それぞれ「報告された成功率」と「真の成功率」を記録する。

6.8 評価

先の実験の結果、「報告された成功率」と「真の成功率」の両方が 100%になった場合を「不正防止の成功」とし、誠実なエージェント (タイプ A) の数と「不正防止の成功」に至った割合をプロットしたものが、図 6.2 である。(エージェント数 8 の場合は、エージェントの組み合わせが 1 通りしか存在しないため、個別に試行を行い結果を集計している。) 誠実なエージェントの数が 0 体の場合であっても不正が防止される構成が存在し、6 体以上の場合はサンプリングした全ての構成で不正の防止が成功していた。

6.9 結論

実験とその評価から、「告発」という限定合理性を仮定した上で「評判スコア」の変化量を決定することで、成員の構成によっては「約束・評判ゲーム」において不正を防止することができるとわかった。これによって、成員の行動を観察できない外部の強制執行力が存在すると仮定した上で、集団を構成する成員達の性質によっては、成員達に合意された約束を履行させられることがわかった。また、成員の構成と不正防止の成功成功の関係性については、先の実験の図 6.2 のとおりである。

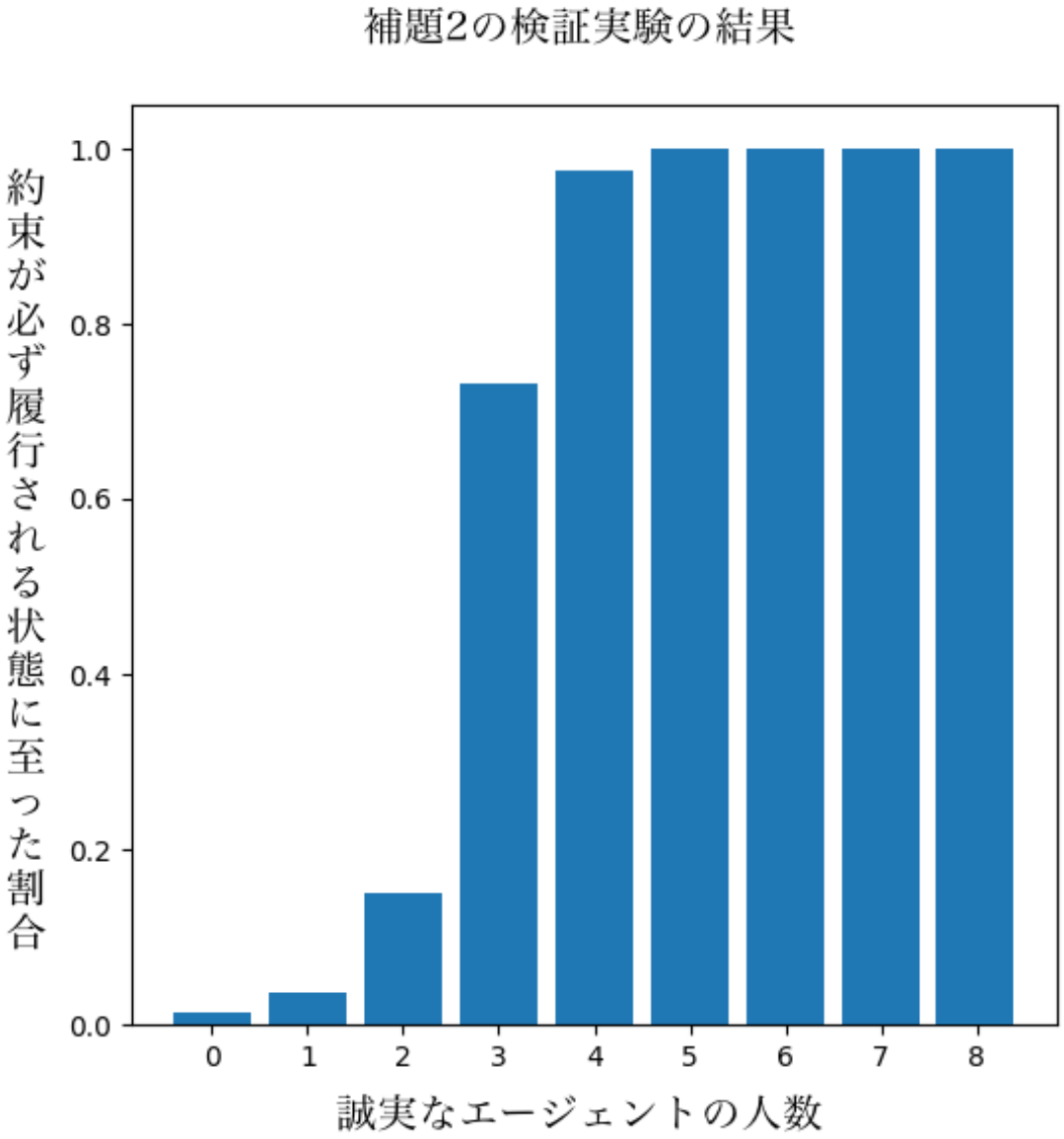

図 6.2: 誠実なエージェントの数と「不正防止の成功」に至った割合

第7章　補題3の検証

本章では、4.1 節で定義した補題 3 を検証する。

7.1　補題3

成員の振る舞いによって補題 2 の強制執行力を自己組織化することができる。

7.2　提案手法

各成員が「評判システム」を 1 つずつ所持している仮定する。補題 1・2 の検証で外部の強制執行力として存在していた「評判システム」をビザンチン将軍問題の署名付きの解法を用いて「分散型評判システム」として設計する。この「評判システム」を所持している成員達のふるまいを記述し、検証 2 の結果と同様の結果になることをマルチエージェントシミュレーションを用いて検証する。

7.3　分散型評判システム

「評判システム」は、各成員の初期の「評判スコア」と報告された各時刻の約束の記録から現時刻の各成員の「評判スコア」を決定するシステムである。各成員が「評判システム」を所有している場合、初期の「評判スコア」と各時刻の約束の記録が互いに共有されていれば、各成員の所有する「評判システム」の現時点の各成員の「評判スコア」は一致する。

7.3.1　仮定

仮定 1 送信されたすべてのメッセージは正しく到達する

仮定 2 メッセージの受信者は誰が送信したのかわかる

仮定 3 メッセージが届かないことを検知できる

仮定 4 誠実な成員の署名は偽造できず、署名されたメッセージの内容が変更されても、それを検知することができる。

仮定 5 誰でも成員の署名の信憑性を検証することができる。

7.3.2 成員の振る舞い

- 事前の合意に基づいて、全ての成員の人数 n を決定する。
- 事前の合意に基づいて、各成員の初期の「評判スコア」$b_i, ..., b_n$ を定義する。
- 事前の合意に基づいて、全ての成員が *promisor* と *reporter* のそれぞれの役割で総当りする周期を定義する。(周期の長さは $n * (n-1)$)
- 事前の合意に基づいて、「評判システム」を用意する。
- 事前の合意に基づいて、約束の内容を定義する。
- 各時刻 t において、「時刻 t における成員の振る舞い」を上から順に実行する。

時刻 t における成員の振る舞い

1. 事前に合意した周期から *promisor* と *reporter* を決定する。
2. *reporter* は時刻 $t - n(n-1)$ に *promisor* と交わした約束の結果を決定する。$t \leq n(n-1)$ の場合、「成功」とする。
3. $V_i = \emptyset$ として初期化する。($\emptyset$ は空集合)
4. *reporter* は約束の結果を全ての成員に署名して送る。
5. 各 i について、
   (a) もし $member_i$ が $v:0$ という形式のメッセージを受け取り、まだ何の報告も受けていない場合は、
      i. $member_i$ は V_i を v にする。
      ii. $member_i$ は他のすべての成員にメッセージ $v:0:i$ を送ります。
   (b) もし $member_i$ が $v:0:j_1:...:j_k$ という形式のメッセージを受け取り、v が集合 V_i に入っていない場合は
      i. $member_i$ は v を V_i に追加する。
      ii. もし $k < m$ であれば、$member_i$ は $j_1, ..., j_k$ 以外のすべての副官に $v:0:j_1:...:j_k:i$ というメッセージを送る。
6. 各 i について、$member_i$ がこれ以上メッセージを受け取らない場合、$member_i$ は $choice(V_i)$ を報告された結果とする。
7. 各 i について、$member_i$ は報告された結果を自身の「評判システム」に記録する。
8. *reporter* は *promisor* と新たな約束を交わす。

7.4 実験方法

この「8 種類のエージェント」から重複問わずランダムに 8 体のエージェントを用意し試行を実施する。これをタイプ A のエージェントが 0〜7 体を占める場合について、それぞれ 8000 回の試行が行われるまで繰り返し、「報告された成功率」と「真の成功率」を記録する。

7.4.1 実験用の時刻 t における成員の振る舞い

1. 事前に合意した周期から *promisor* と *reporter* を決定する。
2. *reporter* は時刻 $t - n(n-1)$ に *promisor* と交わした約束の結果を決定する。$t \leq n(n-1)$ の場合、「成功」とする。
3. $V_i = \emptyset$ として初期化する。($\emptyset$ は空集合)
4. *reporter* は約束の結果を全ての成員に署名して送る。
5. 各 i について、
   (a) もし $member_i$ が $v:0$ という形式のメッセージを受け取り、まだ何の報告も受けていない場合は、
      i. $member_i$ は V_i を v にする。
      ii. $member_i$ は他のすべての成員にメッセージ $v:0:i$ を送ります。
   (b) もし $member_i$ が $v:0:j_1:...:j_k$ という形式のメッセージを受け取り、v が集合 V_i に入っていない場合は
      i. $member_i$ は v を V_i に追加する。
      ii. もし $k < m$ であれば、$member_i$ は $j_1, ..., j_k$ 以外のすべての副官に $v:0:j_1:...:j_k:i$ というメッセージを送る。
6. 各 i について、$member_i$ がこれ以上メッセージを受け取らない場合、$member_i$ は $choice(V_i)$ を報告された結果とする。
7. 各 i について、$member_i$ は報告された結果を自身の「評判システム」に記録する。
8. 各 i について、TODO:上のビザンチン将軍問題で署名送ってないノードをうまく判別して、次に *promisor* になったときに失敗を報告できるようにする。
9. *reporter* は *promisor* と新たな約束を交わす。

7.4.2 ８種類のエージェント

A〜H の 8 種類のエージェントを用意する。それぞれのエージェントは下記の性質に従う。(記述のない振る舞いについては「実験用の時刻 t における成員の振る舞い」に従う。)

A 完全に「成員の振る舞い」に従うエージェント

B 必ず「成功」を報告するエージェント

C 真の約束のと逆の結果を報告するエージェント

D 必ず「失敗」を報告するエージェント

E step6 でタイプ A にだけ結果を送らないエージェント

F step6 でタイプ A にだけ結果を送らず、必ず「成功」を報告するエージェント

G step6 でタイプ A にだけ結果を送らず、真の約束の結果と逆の結果を報告するエージェント

H step6 でタイプ A にだけ結果を送らず、必ず「失敗」を報告するエージェント

7.4.3 事前の合意内容

1. プレイヤー数 n を 8 とする。
2. 初期の各成員の「評判スコア」$b_i, ..., b_n$ を全て 8 とし、各エージェントはそれに合意する。
3. 各時刻 t の *promisor* と *reporter* 組み合わせの周期を $(1,2),(2,1),...,(7,8)$ とする。(周期の長さは 56)
4. 「評判システム」の詳細は検証 2 のものと同様とする。

7.4.4 試行

1. 現在の時刻 t を 0 とする。
2. 各エージェントは「事前の合意内容」に合意する。
3. 時刻 t を 1 進める。
4. 各エージェントはその特性に則って振る舞う。
5. 各エージェントの「評判システム」について、時刻 t の入力によって誰の通貨保有量も 0 未満にならない場合、「報告された結果」を記録する。
6. 各エージェントの「評判システム」について、時刻 t の入力によって誰の通貨保有量も 0 未満にならない場合、「真の結果」を記録する。
7. 時刻 t が 1120 未満なら、step4 に戻る。
8. 各エージェントの「評判システム」において、過去 56 回分の step5 と 6 で記録された結果を集計し、それぞれ「報告された成功率」と「真の成功率」して記録する。

7.5 評価

サンプリングした結果を元に、タイプ A のエージェントと自己組織化が成功した割合をプロットすると、図 7.1 のようになる。この図から、誠実なプレイヤーが 6 人以上の場合は全てのサンプルで自己組織化に成功していることがわかる。

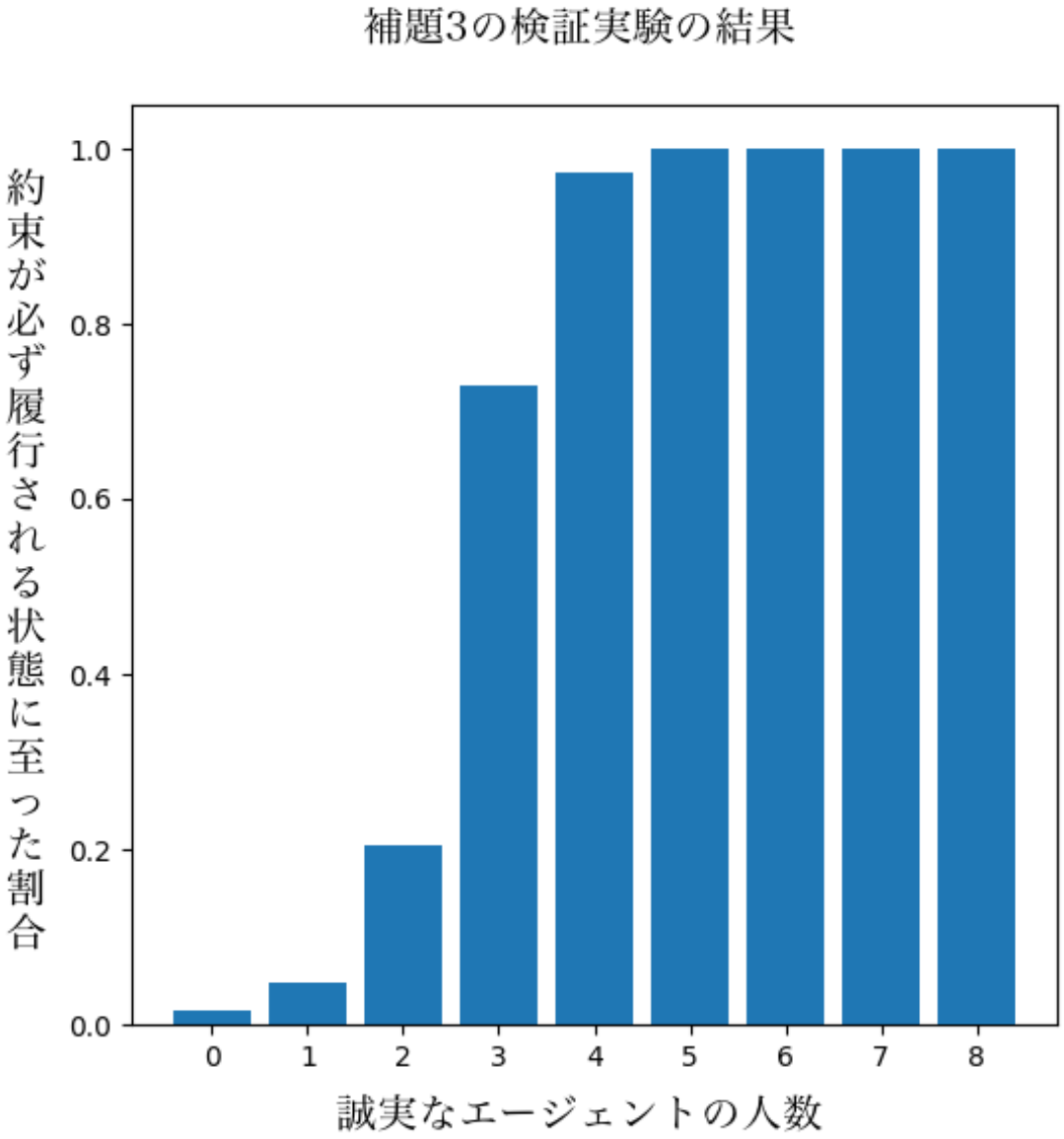

図 7.1: 誠実なエージェントの数と自己組織化に成功した割合

第8章　結論

8.1　本論のまとめ

本研究では、外部の強制執行力が存在しない場合に社会契約が成立しうるかという問いに対して、3 つの補題に分けてそれぞれを検証した。第 1 の補題については、ゲーム理論的なアプローチによって、完全に合理的な成員のみで構成される集団において社会契約が成立しないことを示した。次の補題については、進化ゲーム理論的なアプローチによって、一部の成員が限定合理的であれば社会契約が成立しうることを示した。そして最後の補題については、ビザンチン将軍問題の署名付きの解法を用いることで補題 2 の強制執行力を集団の内部で自己組織化することができることを示した。これらの検証の結果から、下記の仮定の上で一部の成員が誠実に振る舞っていれば社会契約が成立することがわかる。このときの誠実な成員の人数と社会契約の成立の関係性については前章の図 7.1 が示す通りであり、ここから誠実な成員が半数を超える場合には必ず約束が履行される状態に至っていたことがわかる。また、歴史と社会指標、各成員の振る舞いについてはそれぞれ下記のように記すことができる。

8.1.1　仮定

1. 報告者 ($reporter$) 以外の成員は約諾者 ($promisor$) の行動を観察できない。
2. 各成員は同期された時刻を知ることができる。
3. 各成員は集団の人数について事前に合意している。
4. 各成員は成員の社会指標の決定方法について事前に合意している。
5. 各成員は各時刻の約諾者と報告者について事前に合意している。
6. 送信されたすべてのメッセージは正しく到達する
7. メッセージの受信者は誰が送信したのかわかる
8. 各成員はメッセージが届かないことを検知できる
9. 誠実な成員の署名は偽造できず、署名されたメッセージの内容が変更されても、それを検知することができる。
10. 誰でも成員の署名の信憑性を検証することができる。

8.1.2　歴史

社会契約に最低限必要となる歴史とは、各時刻 t における約束の記録の集合である。約束の記録とは報告された時刻、約諾者 ($promisor$)、報告者 ($reporter$)、約束の結果 (「履行」もしくは「反故」) の 4 つの情報からなるものである。

8.1.3 社会指標

社会指標とは、初期の評判スコアと過去の歴史から導き出される評判の指標である。6.5.9 節で導き出した条件を満たすとき、告発する成員が一定数以上いれば約束を保護にする成員の社会指標を下げることができる。

8.1.4 成員の振る舞い

- 事前の合意に基づいて、全ての成員の人数 n を決定する。
- 事前の合意に基づいて、各成員の初期の「評判スコア」$b_i, ..., b_n$ を定義する。
- 事前の合意に基づいて、全ての成員が *promisor* と *reporter* のそれぞれの役割で総当りする周期を定義する。(周期の長さは $n * (n-1)$)
- 事前の合意に基づいて、「評判システム」を用意する。
- 事前の合意に基づいて、約束の内容を定義する。
- 各時刻 t において、「時刻 t における成員の振る舞い」を上から順に実行する。

時刻 t における成員の振る舞い

1. 事前に合意した周期から *promisor* と *reporter* を決定する。
2. *reporter* は時刻 $t - n(n-1)$ に *promisor* と交わした約束の結果を決定する。$t \leq n(n-1)$ の場合、「成功」とする。
3. $V_i = \emptyset$ として初期化する。($\emptyset$ は空集合)
4. *reporter* は約束の結果を全ての成員に署名して送る。
5. 各 i について、
   (a) もし $member_i$ が $v : 0$ という形式のメッセージを受け取り、まだ何の報告も受けていない場合は、
      i. $member_i$ は V_i を v にする。
      ii. $member_i$ は他のすべての成員にメッセージ $v : 0 : i$ を送ります。
   (b) もし $member_i$ が $v : 0 : j_1 : ... : j_k$ という形式のメッセージを受け取り、v が集合 V_i に入っていない場合は
      i. $member_i$ は v を V_i に追加する。
      ii. もし $k < m$ であれば、$member_i$ は $j_1, ..., j_k$ 以外のすべての副官に $v : 0 : j_1 : ... : j_k : i$ というメッセージを送る。
6. 各 i について、$member_i$ がこれ以上メッセージを受け取らない場合、$member_i$ は $choice(V_i)$ を報告された結果とする。
7. 各 i について、$member_i$ は報告された結果を自身の「評判システム」に記録する。
8. *reporter* は *promisor* と新たな約束を交わす。

8.2 本研究の課題

本研究には次に挙げる 3 つの課題がある。

第 1 に、導き出した誠実な成員と社会契約の成立の関係性はヒューリスティックな解にすぎない点が挙げられる。本研究の補題 2 と 3 の検証のために扱った実験は、8 種類のエージェント 8 体の構成は順序も含めて $16,777,216(=8^8)$ 通り考えられるうちの $64,000$ 件のサンプルから導き出されるものに過ぎない。そのため誠実な成員が過半数以上いる場合に社会契約が成立するのかについては慎重な議論が必要である。仮に厳密な誠実な成員と社会契約の成立の関係性を知ろうとするならば、全ての可能性を総当りする大規模な実験を行うか、理論的な解明が必要となるだろう。

第 2 に、補題 3 の検証で用いたビザンチン将軍問題の解法が現実的でない点が挙げられる。本研究では補題 2 と補題 3 の検証結果の類似性を示すために、成員の人数が $f+1$(f は誠実でない成員の人数) の場合にビザンチンフォールトトレラントの達成されるビザンチン将軍問題の初期の署名付きの解法 [15] を用いたが、この解法は同期的なシステムを前提としており、非同期システムにおいては成立しないことが知られている。[9] そのため、より現実的な想定をするならば、pBFT のようなアルゴリズムを用いて成員の振る舞いを記述するべきである。[6] こちらは成員の人数が $3f+1$ 以上の場合にビザンチンフォールトトレラントが達成されるため、実験のデザインには注意が必要となるだろう。

第 3 に、成員の人数が時間経過とともに変動する場合についてモデルを拡張する必要がある。現実の社会に目を向けたとき集団の人数は時間とともに変化していくものであるが、本研究のモデルはこうした集団の人数の変化を考慮されていない。これを実現するためには、成員の人数や各時刻の約諾者と報告者の組み合わせなど、いくつかの事前の合意事項を更新する必要がある。これには分散システムにおける Sybil Attack[7] のように、複数の架空の成員を追加することで社会契約の成立を妨げる攻撃が想定されるため、追加される成員の評判スコアや最低信頼度について慎重に議論する必要がある。

これらは「本研究の動機 (1.1 節)」で述べたような社会契約のインターネットへの適用という目的のためには避けては通れない課題である。

8.3 今後の研究

今後の研究としてはこれら 3 つの課題の解決が必要とされるが、とりわけはじめの 2 つの課題が重要であると考える。仮にその 2 つの課題が解決されるならば、時間経過とともに f が増加する場合に $3f+1$ 以下の人数でもビザンチンフォールトトレラントを達成できる可能性がある。もし誠実でない誠意の数 f に対して成員が $2f+1$ 以上存在する場合に約束が必ず履行される状態に至り、pBFT[6] のような $3f+1$ 以上の成員を必要とするアルゴリズムでも本研究と同じ結果が得られるのであれば、成員の人数が $3f+1$ 以上のとき、本研究の補題 3 の検証結果と同様に、ビザンチン故障の原因になりうる成員が排除できるためである。

例えば 8 人中 2 人の成員が誠実でない場合、$3f+1$(と同時に $2f+1$) を満たしているため、時間経過とともに誠実でない 2 人の成員は排除され、必ず pBFT のアルゴリズムが守られる状態に至る。その後、誠実だった 6 人の成員のうち 1 人が不誠実になったとしても、これは $3f+1$ を満たしているため、ビザンチンフォールトトレラントが達成されて、その 1 人も排除されることになる。結果的に見れば 8 人中 3 人が不誠実な成員という $3f+1$ に反する状況下でもビザンチンフォールトトレラントを達成することができることとなる。

無論、検証するまでこれが確かであるかはわからないが、これが可能であるならば第 3 の課題のような新規に成員が増える場合についても何らかの解決の糸口になりそうである。

また、今回用いたビザンチン将軍問題の解法をはじめとした計算機科学における分散システムの知見は社会契約のモデリングに大いに役立つと考えられる。(EigenTrust[14] などは非常に親しい概念を取り扱っ

ている。) また、逆にこれまで社会契約の議論の中で用いられてきた様々なアプローチは、計算機科学における分散システムの研究に何らかの形で影響を与えるかもしれない。そういった今後この両分野の関係性について、より詳細な研究が必要となるだろう。

謝辞

本研究を進めるにあたり、2016 年度秋学期から所属していた研究軍団 NECO 及び RG(村井・徳田・中村・楠本・高汐・バンミーター・植原・三次・中澤・手塚・武田合同研究会) の皆様、慶應 SFC の教員の皆様、大学進学を支援してくださった家族に感謝します。中でも早稲田大学 大学院経営管理研究科 教授 斉藤賢爾氏には大変お世話になりました。本研究が全く問題意識をはっきりとしない頃から議論にお付き合いいただいたことは勿論、斉藤氏が著書「信用の新世紀 ブロックチェーン後の未来」[20] の中で描かれるこの社会の未来像はこの研究の大きなモチベーションになりました。彼のデジタルマネーに関する思想に出会わなければ、本研究における問題意識を抱くこともなく社会契約を研究対象として選ぶこともなかったと言えます。また、代々 KGL として NECO の運営を担ってくださった阿部涼介氏、菅藤 佑太氏、島津翔太氏、渡邉 聡紀氏にも感謝いたします。特に慶應義塾大学 政策・メディア研究科 博士課程の阿部涼介氏は自分にも他人にも厳しい方でしたが、加えて、彼の研究への姿勢から多くを学ぶことができました。日頃より、RG の運営や講義、研究発表でお世話になっている慶應義塾大学教授 村井純博士、同学部教授 中村修博士、同学部教授 楠本博之博士、同学部教授 高汐一紀博士、同学部教授 Rodney D. Van Meter III 博士、同学部准教授 植原啓介博士、同学部教授 三次仁博士、同学部教授 中澤仁博士、同学部教授 武田圭史博士、同大学政策・メディア研究 科特任准教授 佐藤雅明博士、同大学政策・メディア研究科特任教授 鈴木茂哉博士に感謝いたします。そうした運営に携わってくださった RG Coordinator の方々もありがとうございます。最後に、慶應 SFC という素敵なキャンパスでの大学生活を支えてくださった教員の皆さまや友人、家族に感謝しております。

参考文献

[1] Facebook. https://facebook.com/.

[2] Google. https://google.com/.

[3] Ken Binmore. *Playing Fair (Game Theory and the Social Contract; vol 1)*. Cambridge: The MIT Press, 1994.

[4] Ken Binmore. *Just Playing (Game Theory and the Social Contract, vol. 2)*. Cambridge: The MIT Press, 1998.

[5] Ken Binmore. *Natural Justice*. New York: Oxford University Press, 2005.

[6] Miguel Castro, Barbara Liskov, et al. Practical byzantine fault tolerance. In *OSDI*, Vol. 99, pp. 173–186, 1999.

[7] John (JD) Douceur. The sybil attack. In *Proceedings of 1st International Workshop on Peer-to-Peer Systems (IPTPS)*, January 2002.

[8] Ullmann-Margalit Edna. The emergence of norms, 1977.

[9] Michael J Fischer, Nancy A Lynch, and Michael S Paterson. Impossibility of distributed consensus with one faulty process. *Journal of the ACM (JACM)*, Vol. 32, No. 2, pp. 374–382, 1985.

[10] David Gauthier. *Morals by agreement*. Oxford University Press on Demand, 1986.

[11] GSMA. The mobile economy 2020, 2020.

[12] WernerGüth, Rolf Schmittberger, Bernd Schwarze. An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior & Organization*, Vol. 3, No. 4, pp. 367–388, 1982.

[13] John C Harsanyi. Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility. *Journal of Political Economy*, Vol. 63, No. 4, pp. 309–321, 1955.

[14] Sepandar D Kamvar, Mario T Schlosser, and Hector Garcia-Molina. The eigentrust algorithm for reputation management in p2p networks. In *Proceedings of the 12th international conference on World Wide Web*, pp. 640–651, 2003.

[15] Leslie Lamport, Robert Shostck, and Marshall Pease. The byzantine generals problem. *ACM Transactions on Programming Languages and Systems*, Vol. 4, No. 3, pp. 382–401, 1982.

[16] John Mackie. *Ethics*. Penguin Books Ltd, 1977.

[17] Satoshi Nakamoto. Bitcoin: A peer-to-peer electronic cash system. 2008.

[18] John Rawls. *A Theory of Justice*. Cambridge: Harvard University Press, 1971.

[19] Jean-Jacques Rousseau. *Discourse on the Origin of Inequality*. 1755.

[20] Kenji Saito. 信用の新世紀　ブロックチェーン後の未来. インプレス R&D, 2017.

[21] Kenji Saito and Shugo Ikemoto. 地球規模 os 外殻 (シェル) の開発と応用. 2008.

[22] Herbert A Simon. *Administrative behavior; a study of decision-making processes in administrative organization*. 1947.

[23] Brian Skyrms. *Evolution of the social contract*. Cambridge University Press, 1996.

[24] Brian Skyrms. The stag hunt. In *Proceedings and Addresses of the American Philosophical Association*, Vol. 75, pp. 31–41. JSTOR, 2001.

[25] Brian Skyrms. *The stag hunt and the evolution of social structure*. Cambridge University Press, 2004.

[26] Robert Sugden. *The Economics of Rights, Co-operation and Welfare*. Oxford: Basil Blackwell, 1986.

[27] Bruno Verbeek and Christopher Morris. Game Theory and Ethics. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Winter 2020 edition, 2020.

[28] Gregory Wheeler. Bounded Rationality. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Fall 2020 edition, 2020.