

卒業論文 2021 年度 (令和 3 年)

「商取引ゲーム」をベースにした  
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 - 2020 年度 (令和 2 年度)

「商取引ゲーム」をベースにした  
計算可能な複雑系としての社会契約

キーワード:

1. 社会契約, 2. 複雑系, 3. マルチエージェントシミュレーション, 4.

慶應義塾大学 環境情報学部  
宮元 眺

# 目次

<b>第1章 序論</b>	<b>5</b>
1.1 なぜ私たちは法に従うのか？	5
1.2 これまでの社会契約の議論	5
1.3 インターネットと社会契約	6
1.4 社会契約とは	6
1.5 本論の目的	6
1.6 本論のアプローチ	6
1.7 本論の構成	7
<b>第2章 背景</b>	<b>8</b>
2.1 社会契約	8
2.2 デイビッドヒュームのコンヴェンション	8
2.3 スタグハントゲーム	8
2.4 フォーク定理	8
2.5 強制執行力	8
2.6 外部の強制執行力	8
2.7 限定合理性	8
2.8 複雑系	8
2.9 自己組織化	8
2.10 ビザンチン将軍問題	8
2.11	8
<b>第3章 問題定義</b>	<b>9</b>
3.1 本論における問題定義	9
3.2 本論における問題解決の要件	9
3.2.1 歴史を定義する	9
3.2.2 社会指標の計算方法を定義する	9
3.2.3 成員の振る舞いのみで記述できる	9
<b>第4章 仮説と検証方法</b>	<b>10</b>
4.1 本論における仮説	10
4.2 仮説の検証方法	10
<b>第5章 商取引ゲーム</b>	<b>12</b>
5.1 本章の問題提起	12
5.2 本章の仮説	12
5.3 提案手法	12
5.4 行動観察不可の条件	12

5.5	無法地帯での商取引	12
5.5.1	商取引システム	12
5.6	商取引ゲーム	12
5.6.1	ゲームの進め方	13
5.6.2	ゲーム木	13
5.6.3	非協力戦略型ゲーム	14
5.7	不正防止の不可能性の証明	14
5.7.1	前提の整理	14
5.7.2	不正行為が起きない戦略組とその条件	15
5.7.3	命題	15
5.7.4	証明	15
5.8	結論	19
<b>第 6 章</b>	<b>倫理ある商取引ゲーム</b>	<b>20</b>
6.1	本章における問題提起	20
6.2	本章の仮説	20
6.3	提案手法	20
6.3.1	倫理ある商取引ゲーム	20
6.3.2	誠実な戦略をとった割合と成功が報告される割合の関係	21
6.4	「倫理ある商取引ゲーム」において不正が防止される条件	21
6.4.1	不正が抑制される戦略組と期待利得の不等式	21
6.4.2	<i>seller</i> と <i>buyer</i> の期待利得	21
6.4.3	<i>seller</i> が誠実な戦略をとる条件	22
6.4.4	<i>buyer</i> が誠実な戦略をとる条件	22
6.4.5	信頼度 $p^{player}$	22
6.4.6	最低信頼度 $T^{player}$	23
6.4.7	最低信頼度を用いた条件	23
6.5	実験方法	23
6.5.1	エージェントの種類	23
6.5.2	試行	24
6.6	「商取引システム」の詳細	25
6.6.1	ReputationWeight	25
6.6.2	「成功」が報告された場合の通貨保有量の変化	25
6.6.3	EscrowCost	25
6.6.4	EscrowCost の負担比率	25
6.6.5	EscrowCost の分配	26
6.7	評価	26
6.8	結論	26
<b>第 7 章</b>	<b>複雑系としての社会契約</b>	<b>29</b>
7.1	本章で取り組む問題	29
7.2	本章で検証する仮説	29
7.3	提案手法	29
7.3.1	「商取引システム」を自己組織化する商取引契約	30

7.4	実験方法 . . . . .	30
7.4.1	実験用の商取引契約 . . . . .	31
7.4.2	8 種類のエージェント . . . . .	31
7.5	評価 . . . . .	31
<b>第 8 章</b>	<b>結論</b>	<b>33</b>
8.1	本論のまとめ . . . . .	33
8.2	本論の課題 . . . . .	33
8.3	今後の研究 . . . . .	33

# 第1章 序論

## 1.1 なぜ私たちは法に従うのか？

普段、我々は、人々が法に従うことを当然のように信じており、この問いについて深く考えることはめったにない。例えば、あなたが近所のお店でコーヒーを注文したとしよう。あなたはメニューを見て、注文を行い、代金を支払って、コーヒーを受け取る。結果的に、あなたはコーヒーを手にして、店主は代金を手にする。日常的に経験するであろう法の守られた世界である。

しかし、仮に、あなたがメニューを見て、注文を行い、請求された金額を支払ったのに、店主が飲み物を渡さなかったとしよう。あなたは次にどうするだろうか。警察に通報するだろうか、裁判所に訴えて返金を要求するだろうか。もし警察に通報したり、裁判所に訴えたりしたとして、彼らと店主が共謀していて、不正をなかったことにされたらどうするだろうか。

我々は無意識にどこかで、店主を信頼している。そうでなくとも、店主が不正を働いた場合に、警察が解決してくれることを信頼している。仮に店主と警察の間で賄賂が渡されており、両者が共謀していたとしても、裁判所が正常に機能していることを信頼している。裁判所が機能しなくても不正を行ったものが他の何かを罰してくれることを信じている。

つまるところ、我々は、不正が起きても当事者でない強制力をもった機関 (外部の強制執行機関) がその不正を正してくれることを信じている。そうでなくとも、その機関が不正を働いても、さらなる外部の強制執行機関がその不正を正してくることを信じている。そして、我々は暗黙のうちにどこかでこの信頼の連鎖の根源を考えることをやめてしまう。

なぜ店主は信頼できるのか？なぜ警察は信頼できるのか？なぜ裁判所は信頼できるのか？なぜ国家は信頼できるのか？なぜ法律は信頼できるのか？なぜ立法機関は信頼できるのか？なぜ国会議員は信頼できるのか？なぜ国民は信頼できるのか？

この信頼の連鎖の根源を探ることが、「なぜ私たちは法に従うのか？」という問いに、理論的な答えを導き出すことに繋がるだろう。

## 1.2 これまでの社会契約の議論

これまで、この問いは社会契約という領域で研究されてきた。近代において、はホブズやロック、ルソーによって、民主国家の正当性を示すために議論され、王権神授説に対する反論としてフランス革命の理論的な石杖となった。現代においては、ロールズやハースニーが彼らの社会契約を一般化したモデルを構築し、「公正な正義とはいかなるものか？」という問いへの議論に発展した。

社会契約の議論は伝統的に、自然状態と社会状態の2つの状態を仮定することから始まる。自然状態とは、ある集団において、その成員が法に従うことに強制力の働かない状態であり、社会状態とは、全ての成員が法に従うことに強制力の働く状態である。この自然状態から社会状態へと遷移するプロセスを明らかにすることで、法に従うことに強制力が働くメカニズムの解明を試みてきた。

最近では、このメカニズムをゲーム理論や進化ゲーム理論を用いて解析した研究も増えており、「道徳とは何か？」、「どのような場合に我々は法に従うのか？」、「どういったプロセスで法に従うのか」というよ

り派生した問いに主眼が置かれている。

### 1.3 インターネットと社会契約

社会の動向の変化とともに、社会契約のメカニズムの解明には、新たな意義が生まれてきている。

その最もたる変化は、インターネットが国家を超越した地球のインフラになろうとしていることである。今や、Google や Facebook のようなサービスは世界中にユーザーを抱えており、提供元となっている企業は複数の国家にその支社を構える多国籍企業となっている。こうした企業のガバナンスはもはや一国の法にのみ依存するものではない。

また、近年では、「誰にも送金を止めさせない」という目的に基づいた Bitcoin のようなブロックチェーン技術の登場により、国家に依存しない形でシステムを運用しようとする試みが増えている。こうした動きは「地球規模 OS」のような概念の必要性を感じさせる。

こうしたインターネットの急速なグローバル化から生じるのは、国家を超越したインターネット上のサービスの正当性を誰が担保するのかという問題である。これまではサービスの正当性は、国家の法に縛られた企業が保証していればよかったが、これからはインターネットを国家を超越した地球のインフラと考え、そのガバナンスを担保する新たな取り決めが国家を超えたところで必要となる。それゆえ、国家を超越したインターネット上のサービスの正当性を保証するためにも、そのサービスのユーザーがコンピューターを用いて社会契約を果たせるモデルの構築が必要である。

### 1.4 社会契約とは

ここまでの議論から、本論における社会契約とは、ある集団において、その一員もしくは全員がある取り決めに合意することによって、その取り決めに強制力を生じさせるプロセスと定義する。

### 1.5 本論の目的

本論の目的は社会契約を一般化した計算可能なモデルを構築し、コンピューターシミュレーションを用いて、ある集団の成員の性質と社会契約の成立の関係性を明らかにすることである。

先に述べたように、我々の見据えるインターネット上での社会契約において、人々はコンピューターを用いて、それを成すことになる。ならば社会契約のモデルは計算可能なレベルで一般化される必要があるだろう。また、はじめに述べたように、我々の生きる社会は完全な無法地帯でなければ、全員が法に従っているわけでもない。中には、法に反した行動をとりながらも、モデルが完成しただけでは実用にはいたれない、そのシステムがいかなる場合にどれだけ安全であるのかを知る必要がある。そのため、モデルを構築するとともに、コンピューターシミュレーションによってその安全性を検証する。

### 1.6 本論のアプローチ

本論では、ある集団において外部の強制執行力に依らずに取り決めが遵守されるのは、各成員の振る舞いによる複雑系の中で取り決めを遵守させるシステムが自己組織化されているためだと考え、その複雑系の詳細を商取引ゲームというアイディアを基にモデリングを試みる。商取引ゲームとは、商取引システムという唯一の外部の強制執行力の存在を仮定した上で、商取引をモデリングした非協力戦略型ゲームである。この商取引システムは、商取引契約が履行されたか否かの真の結果を観察することはできないが、プレイ

ヤーから報告された結果を観察することができ、それに応じて各プレイヤーの通貨の保有量を操作することのできるシステムである。

本論のアプローチは主に 2 つの段階に分けることができる。1 つは、この商取引システムの通貨の保有量を操作する機能を用いて、商取引ゲームにおいて不正が生じないようなインセンティブ設計を行うことであり、もう 1 つは、商取引契約の内容にプレイヤーの振る舞いを記述することで、この商取引システムを複雑系の内部で自己組織化されたシステムとして設計することである。

この 2 つのアプローチが成功したとき、商取引契約の内容には、不正を防止可能なインセンティブ設計を行える商取引システムを～が定義されており、この商取引契約はその契約によって自己組織化された内部の商取引システムによって履行されるようになるはずである。

ここで商取引契約の内容とは先に定義した社会契約における取り決めであり、プレイヤーとは同様にある集団の成員であるため、ここで本論の目的である、社会契約を一般化した計算可能なモデルの構築が果たされることとなる。

一見、契約を履行させるシステムによってそのシステムの正当性を担保する契約を履行させるという論理は、循環をはらんでいるように思えるだろうが、実際のところはこの 2 つの時間軸に依存しあっている。それ故に初期値によってその社会契約が成立するか否かは全く異なるものとなり、その結果を全て把握することは困難である。

しかしながら、幸いなことにこの複雑系としての社会契約は計算可能なモデルであるため、そのカオスをコンピューターシミュレーションによって近似することができるだろう。

原点に立ち返るなら「なぜ我々は法を守るのか」という疑問に対して、本論が提示できる答えは「他の人も法を守っているから」というものになるだろう。では具体的にどのような性質の成員がどのくらいいれば法が守られるのか、それが本論の提示する複雑系としての社会契約のモデルから明らかになるだろう。

## 1.7 本論の構成

先に上げた 2 つのアプローチについて、前者の不正を防止する商取引システムのインセンティブ設計は 3 章と 4 章で、後者の商取引システムを自己組織化する商取引契約の内容の設計は 5 章で説明する。

3 章では利己的なプレイヤーのみで行われる商取引ゲームにおいて、報告された結果に基づいて通貨の保有量を操作しても、不正を防止することは不可能であることを示す。

4 章では、商取引に失敗した場合にその結果を正しくシステムに報告するプレイヤーを倫理あるプレイヤーと考し、こうしたプレイヤーのみで行われる「倫理ある商取引ゲーム」について考える。この限定合理性の上でインセンティブ設計を行うことで、倫理あるプレイヤー以外のプレイヤーが参加する商取引ゲームであっても、参加するプレイヤー達の性質によっては不正を防止することが可能であることを示す。

5 章では、商取引システムを運営する約束を、商取引ゲームで取引されていた商品 (サービス) とすることで、外部の強制執行力として存在していた商取引システムを自己組織化させる。これによって、外部の強制執行力に依存することなく、社会の成員たちによって構成された複雑系の内部で商取引システムが稼働し、「倫理ある商取引ゲーム」と同様に参加するプレイヤーの性質によって不正が起きなくなることを示す。

6 章では、本論では、ここまでの内容をまとめるとともに、今回紹介した複雑系としての社会契約のモデルの課題を上げ、今後の研究についての計画述べる。



## 第2章 背景

2.1 社会契約

2.2 デイビッドヒュームのコンヴェンション

2.3 スタグハントゲーム

2.4 フォーク定理

2.5 強制執行力

2.6 外部の強制執行力

2.7 限定合理性

2.8 複雑系

2.9 自己組織化

2.10 ビザンチン将軍問題

2.11

## 第3章 問題定義

### 3.1 本論における問題定義

これまでのゲーム理論を用いて社会契約のメカニズムを説明しようとする研究では、強制執行力の存在が仮定されている、もしくは外部の強制執行力の存在が仮定されている上で議論が進められていた。外部の強制執行力が存在している場合については、別の社会契約が成立している必要があるため、その別の社会契約が成立しているメカニズムを説明する必要性が生じ問題が堂々巡りに陥る。一方、外部の強制執行力が存在しない場合については、社会の過去の歴史が決定する社会指標を利用することで社会契約が成立するとされている (Binmore 2005[5]) が、この研究は還元主義的であり具体的に成員達がどのような振る舞いをするかで、そうした歴史が決定されどのように社会指標を導き出すかを決定することは困難であった。

そこで本研究では、改めて外部の強制執行力が存在しない場合について、合意された約束を成員達に遵守させることが可能なのかという問題に取り組み、成員達のどのような振る舞いによって、歴史を決定し、社会指標を導き出せばよいのかを計算可能なレベルで明確にする。

### 3.2 本論における問題解決の要件

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を作り出せていることを示す他に、下記の3つの要件が満たされている必要がある。

#### 3.2.1 歴史を定義する

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 社会指標の計算方法を定義する

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 成員の振る舞いのみで記述できる

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 本論における仮説

外部の強制執行力が存在しない場合であっても、集団を構成する成員の性質によっては強制執行力を生み出すことが可能であり、それによって成員たちに合意した約束を遵守させることが可能である。

### 4.2 仮説の検証方法

本論では、3つの補題の仮説検証を行うことで、先の仮説を示す。

第1に、全ての成員が完全に合理的である場合、成員の振る舞いによって決定された歴史から各成員の社会指標を算出できる外部の強制執行力が、成員達に約束を遵守させるようなインセンティブ設計をできないことを示す。私達は、この検証のために、任意の約束を商取引契約として結ぶ売り手と買い手の商取引について考える。具体的には、売り手が報告した商取引の結果を歴史として記録し、各成員の社会指標である通貨保有量を決定することができる外部の強制執行力としての「商取引システム」の存在を仮定する。また、この「商取引システム」を用いて行われる商取引を「商取引ゲーム」という非協力戦略型ゲームとしてモデリングする。そして、買い手と売り手の各戦略の利得を比較することで、全ての成員が合理的である場合に「商取引システム」から報告された商取引の結果に基づいて不正が防止されるような買い手と売り手の通貨保有量を決定することが不可能であることを示す。

第2に、一部あるいは全部の成員が限定合理的である場合、成員の振る舞いによって決定された歴史から各成員の社会指標を算出できる外部の強制執行力が、合理的な成員達に約束を遵守させるようなインセンティブ設計をできることを示す。この検証のために、我々は先の「商取引ゲーム」において報復的な戦略(買い手が商取引契約を履行しなかった場合に「失敗」を報告する戦略)をとる成員のみによって行われる「倫理ある商取引ゲーム」について考える。はじめの検証と同じように、この「倫理ある商取引ゲーム」における買い手と売り手の各戦略の利得を比較することで、このモデルにおいては「商取引システム」から報告された商取引の結果に基づいて不正を防止できるような買い手と売り手の通貨保有量を決定することが可能であることを示す。また、この条件に基づいて各誠意の通貨保有量を操作する商取引システムを設計し、本来の「商取引ゲーム」に適用してマルチエージェントシミュレーションを用いた実験を行うことで、全ての成員が先に述べた報復的な戦略を取る場合でなくても、一部あるいは全部の成員が報復的な戦略をとっていれば、不正が防止されうることを示す。

第3に、外部の強制執行力が存在しない場合でも、一部あるいは全部の成員が限定合理的である場合、各成員の振る舞いのみによって成員達に約束を遵守させる強制執行力を生じさせることが可能であることを示す。この検証のために、我々は先の検証で外部の執行力として存在が仮定されていた「商取引システム」を集団が構成する複雑系の中で各誠意の振る舞いによって自己組織化されたシステムとして設計し、第2の検証と同様の実験を行う。「商取引システム」を各成員の振る舞いによって自己組織化されたシステムとして再現することができ、一部あるいは全部の成員が報復的な戦略をとっていれば、不正が防止されうることを示す。

これら 3 つの補題の検証により、外部の強制執行力が存在しない場合であっても、集団を構成する成員の性質によっては強制執行力を生み出すことが可能であり、それによって成員たちに合意した約束を遵守させることが可能だとわかる。また、問題提起の要件で述べた社会契約の成立に必要な歴史の定義や社会指標の計算方法、具体的な成員の振る舞いがそれぞれ明確になる、

## 第5章 商取引ゲーム

本章では、外部の強制執行力によって商取引契約の履行が保証されない商取引を「商取引ゲーム」としてモデリングし、合理的な集団を仮定した場合に、そのゲームにおいて不正が生じないようなインセンティブを設計が不可能なことを示す。

外部の強制執行力が存在すると仮定した上で、歴史と社会指標の計算方法を定義することで、社会契約が成立するかを確かめるため、

### 5.1 本章の問題提起

通貨を社会指標ととらえ、商取引の結果を合理的な人々が商取引を成功させるようなインセンティブ設計ができるか？

仮にそれをなし得る外部の強制力が存在していたとして、1章の初めに述べたように、その存在というのは堂々巡りになっていく。警察も、裁判所も、税金を支払うことによって受けているサービスという点で、商取引だと考えれば、この商取引の依存関係を紐解いたときに、必ず1つは他の商取引に依存せずとも、契約が履行される商取引が存在しているはずである。果たしてそんな商取引のメカニズムはデザイン可能なのか？

### 5.2 本章の仮説

### 5.3 提案手法

### 5.4 行動観察不可の条件

### 5.5 無法地帯での商取引

この問いに答えるために、本章では無法地帯での商取引について考える。商品(財・サービス)の引き渡しと通貨の支払いについて考える。売り手と買い手(buyer)の2者間で行われる商取引。外部の強制執行力は働かない。

商取引が成功したのか失敗したのかの報告を受けて、それに応じて各プレイヤーの通貨の保有量を変更するシステムである。

#### 5.5.1 商取引システム

### 5.6 商取引ゲーム

また、この「商取引システム」において、商取引で不正行為を防止できるインセンティブ設計を行うことができるのかを検証するために、このシステムを商取引の仲介として用いる「商取引ゲーム」を次

のように定義する．なお，ここでのゲームの *players*(システムの参加者達) は合理的に (期待利得の最も高い) 戦略を決定するものとする．

### 5.6.1 ゲームの進め方

本稿での商取引は，以下の 4 つのステップで *seller* と *buyer* が交互に行動を展開するものとする．

**step1** *seller* は商品 *goods* とその価格を告知する

**step2** その商品の購入を希望する *buyer* が「商取引システム」に商取引の合意を報告する．

**step3** *seller* は「正当な行為」と「不正な行為」のどちらかの行動選択をする．ここで「正当な行為」を行った場合，*buyer* は契約通りの *goods* を受け取れ，「不正な行為」を行った場合，*buyer* は契約通りの *goods* を受け取れないものとする．

**step4** *buyer* は商取引の「成功」か「失敗」かを「商取引システム」に報告する．この報告に基づき，「商取引システム」は *seller* と *buyer* の保有する通貨の量を調整する．

### 5.6.2 ゲーム木

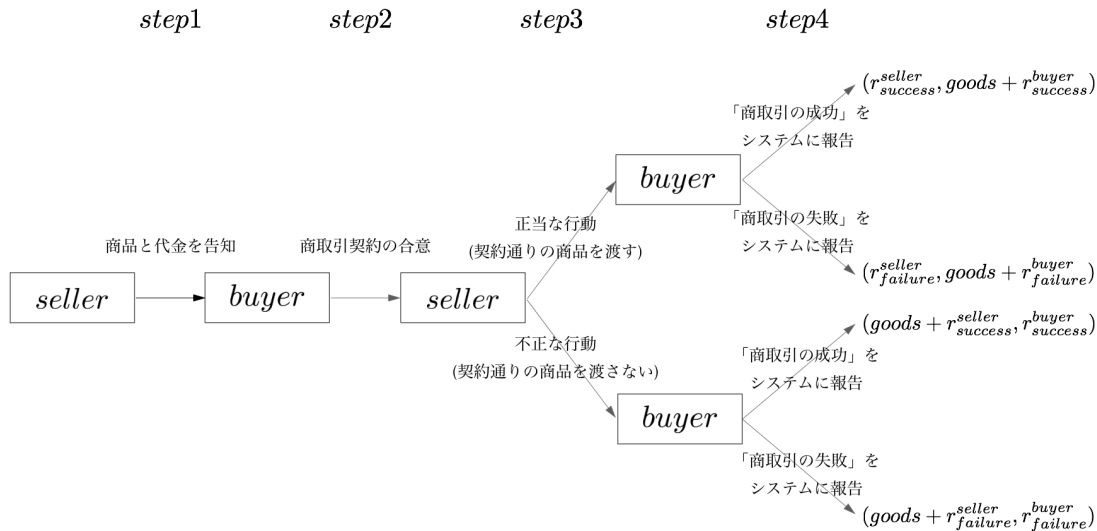


図 5.1: 「商取引ゲーム」のゲーム木

先の商取引ゲームをゲームの木を用いて表すと図 5.1 のようになる．**step1** と **step2** の時点では商取引の結果が変化することはないが，**step3** で *seller* が「正当な行為」をとるか否かと，**step4** で *buyer* からの報告によってのみ商取引の結果は変化する．また，「商取引システム」からは「行動観察不可の条件」より *seller* と *buyer* がどの行動をとったかはわからないため，**step4** で *buyer* の報告にのみ基づいて *seller* と *buyer* の保有する通貨の量を調整しなければならない．つまりは商取引の結果①と③，②と④はそれぞれ *goods* を除く利得は同じでなくてはならない．

### 5.6.3 非協力戦略型ゲーム

		<i>buyer</i>			
		$s_1^{buyer}$	$s_2^{buyer}$	$s_3^{buyer}$	$s_4^{buyer}$
<i>seller</i>	$s_1^{seller}$	$(r_{success}^{seller},$ $goods + r_{success}^{buyer})$	$(r_{success}^{seller},$ $goods + r_{success}^{buyer})$	$(r_{failure}^{seller},$ $goods + r_{failure}^{buyer})$	$(r_{failure}^{seller},$ $goods + r_{failure}^{buyer})$
	$s_2^{seller}$	$(goods + r_{failure}^{seller},$ $r_{failure}^{buyer})$	$(goods + r_{success}^{seller},$ $r_{success}^{buyer})$	$(goods + r_{failure}^{seller},$ $r_{failure}^{buyer})$	$(goods + r_{success}^{seller},$ $r_{success}^{buyer})$

表 5.1: 非協力戦略型ゲームとして表した「商取引ゲーム」の利得表

また、この商取引のモデルは、第 3 ステップ以降の *seller* の行動選択と、それに対する第 4 ステップの *buyer* の行動選択を、非協力戦略型ゲームとしてとらえられる。ここで、戦略  $s_n^{player}$  を *player* (ここでは *seller* か *buyer*) の取りうる戦略番号  $n$  の戦略として、*seller* と *buyer* のそれぞれの戦略は以下のように定義する。

$s_1^{seller}$  ... 正当な行為を行う

$s_2^{seller}$  ... 不正な行為を行う

$s_1^{buyer}$  ... *seller* が正当な行為をとった場合は「成功」を、不正な行動をとった場合は「失敗」を報告する

$s_2^{buyer}$  ... *seller* が正当な行為をとった場合は「成功」を、不正な行動をとった場合は「成功」を報告する

$s_3^{buyer}$  ... *seller* が正当な行為をとった場合は「失敗」を、不正な行動をとった場合は「失敗」を報告する

$s_4^{buyer}$  ... *seller* が正当な行為をとった場合は「失敗」を、不正な行動をとった場合は「成功」を報告する

また、商取引終了時の *player* の保有する通貨量の変化によって生じる利得を、第 4 ステップでの報告が「成功」だった場合は  $r_{success}^{player}$ 、「失敗」だった場合は  $r_{failure}^{player}$  とし、商品の所有によって生じる利得を *goods* と表す。ここで、*seller* と *buyer* の任意の戦略組  $(s_n^{seller}, s_n^{buyer})$  の際の *seller* と *buyer* の各利得は表 5.1 のようになる。

## 5.7 不正防止の不可能性の証明

### 5.7.1 前提の整理

- システムは *buyer* によって報告された商取引の結果を観察できる。
- システムは *seller* と *buyer* がどの戦略を選んだかはわからない。
- 商取引に参加する *player* は合理的に (利得の期待値が最も高い) 戦略を決定する。
- システム内には追跡可能な通貨が存在しており、商取引にはその通貨が用いられる。
- システムからは *player* の通貨の保有量を操作することができる。

### 5.7.2 不正行為が起きない戦略組とその条件

表より、商取引で不正行為が起きないためには、*seller* と *buyer* のとる戦略組が  $(s_1^{seller}, s_1^{buyer})$  もしくは  $(s_1^{seller}, s_2^{buyer})$  のいずれかに帰着しなければならない。

ここで *player* が戦略  $s_n^{player}$  をとったときの利得  $R$  の期待値を  $E(R|s_n^{player})$  とする。

$(s_1^{seller}, s_1^{buyer})$  に帰着するためには、

$$E(R|s_1^{seller}) > E(R|s_2^{seller}) \text{ かつ } E(R|s_1^{buyer}) > E(R|s_2^{buyer}) \text{ かつ } E(R|s_1^{buyer}) > E(R|s_3^{buyer}) \text{ かつ } E(R|s_1^{buyer}) > E(R|s_4^{buyer}) \dots \text{条件①}$$

を満たす必要があり、 $(s_1^{seller}, s_2^{buyer})$  に帰着する場合は、

$$E(R|s_1^{seller}) > E(R|s_2^{seller}) \text{ かつ } E(R|s_2^{buyer}) > E(R|s_1^{buyer}) \text{ かつ } E(R|s_2^{buyer}) > E(R|s_3^{buyer}) \text{ かつ } E(R|s_2^{buyer}) > E(R|s_4^{buyer}) \dots \text{条件②}$$

を満たす必要がある。

つまり、不正行為を防ぐ商取引のインセンティブ設計を行うためには、条件①もしくは条件②を満たす  $(r_{success}^{seller}, r_{failure}^{seller}, r_{success}^{buyer}, r_{failure}^{buyer})$  の組を「商取引システム」から決定することができなければならない。そこで、不正行為を防ぐ商取引のインセンティブ設計が不可能であることを示すために、次の命題を証明する。

### 5.7.3 命題

条件①もしくは条件②のいずれかの条件を満たす  $(r_{success}^{seller}, r_{failure}^{seller}, r_{success}^{buyer}, r_{failure}^{buyer})$  の組を商取引システムから決定することはできない。

### 5.7.4 証明

*player* が戦略  $s_n^{player}$  をとる確率を  $p_n^{player}$  と表す。 ( $0 \leq p_n^{player} \leq 1$ )

*seller* について、各戦略の期待利得は以下のように表せる。

$$E(R|s_1^{seller}) = p_1^{buyer} r_{success}^{seller} + p_2^{buyer} r_{success}^{seller} + p_3^{buyer} r_{failure}^{seller} + p_4^{buyer} r_{failure}^{seller}$$

$$E(R|s_2^{seller}) = p_1^{buyer} (goods + r_{failure}^{seller}) + p_2^{buyer} (goods + r_{success}^{seller}) + p_3^{buyer} (goods + r_{failure}^{seller}) + p_4^{buyer} (goods + r_{success}^{seller})$$

$$= goods + p_1^{buyer} r_{failure}^{seller} + p_2^{buyer} r_{success}^{seller} + p_3^{buyer} r_{failure}^{seller} + p_4^{buyer} r_{success}^{seller}$$

ここで、 $E(R|s_1^{seller}) > E(R|s_2^{seller})$  を満たすためには、

$$p_1^{buyer} r_{success}^{seller} + p_2^{buyer} r_{success}^{seller} + p_3^{buyer} r_{failure}^{seller} + p_4^{buyer} r_{failure}^{seller}$$

$$> goods + p_1^{buyer} r_{failure}^{seller} + p_2^{buyer} r_{success}^{seller} + p_3^{buyer} r_{failure}^{seller} + p_4^{buyer} r_{success}^{seller}$$



$$\therefore p_1^{buyer} r_{success}^{seller} + p_4^{buyer} r_{failure}^{seller} > goods + p_1^{buyer} r_{failure}^{seller} + p_4^{buyer} r_{success}^{seller}$$

$$\therefore p_1^{buyer} r_{success}^{seller} + p_4^{buyer} r_{failure}^{seller} - goods + p_1^{buyer} r_{failure}^{seller} - p_4^{buyer} r_{success}^{seller} > 0$$

$$\therefore p_1^{buyer} (r_{success}^{seller} - r_{failure}^{seller}) - p_4^{buyer} (r_{success}^{seller} - r_{failure}^{seller}) - goods > 0$$

$$\therefore (p_1^{buyer} - p_4^{buyer}) (r_{success}^{seller} - r_{failure}^{seller}) - goods > 0$$

$$\therefore (p_1^{buyer} - p_4^{buyer}) (r_{success}^{seller} - r_{failure}^{seller} - \frac{goods}{p_1^{buyer} - p_4^{buyer}}) > 0$$

を満たす必要がある。つまり,

$$p_1^{buyer} > p_4^{buyer} \text{ のときは,}$$

$$r_{success}^{seller} - r_{failure}^{seller} > \frac{goods}{p_1^{buyer} - p_4^{buyer}}$$

$$p_1^{buyer} < p_4^{buyer} \text{ のときは,}$$

$$r_{success}^{seller} - r_{failure}^{seller} < \frac{goods}{p_1^{buyer} - p_4^{buyer}}$$

を満たせば,  $E(R|s_1^{seller}) > E(R|s_2^{seller})$  である。なお,  $p_1^{buyer} = p_4^{buyer}$  のときは,

$E(R|s_1^{seller}) > E(R|s_2^{seller})$  は成り立たない。

また, *buyer* について, 各戦略の期待利得は以下のように表せる。

$$E(R|s_1^{buyer}) = p_1^{seller} (goods + r_{success}^{buyer}) + p_2^{seller} r_{failure}^{buyer}$$

$$E(R|s_2^{buyer}) = p_1^{seller} (goods + r_{success}^{buyer}) + p_2^{seller} r_{success}^{buyer}$$

$$E(R|s_3^{buyer}) = p_1^{seller} (goods + r_{failure}^{buyer}) + p_2^{seller} r_{failure}^{buyer}$$

$$E(R|s_4^{buyer}) = p_1^{seller} (goods + r_{failure}^{buyer}) + p_2^{seller} r_{success}^{buyer}$$

### 条件①が成り立たないことの証明

ここで, 条件①の必要条件である  $E(R|s_1^{buyer}) > E(R|s_2^{buyer})$  かつ  $E(R|s_1^{buyer}) > E(R|s_3^{buyer})$  かつ  $E(R|s_1^{buyer}) > E(R|s_4^{buyer})$  を満たすためには,

$$E(R|s_1^{buyer}) > E(R|s_2^{buyer})$$

$$\therefore p_1^{seller} (goods + r_{success}^{buyer}) + p_2^{seller} r_{failure}^{buyer} > p_1^{seller} (goods + r_{success}^{buyer}) + p_2^{seller} r_{success}^{buyer}$$

$$\therefore p_2^{\text{seller}} r_{\text{failure}}^{\text{buyer}} - p_2^{\text{seller}} r_{\text{success}}^{\text{buyer}} > 0$$

$$\therefore p_2^{\text{seller}} (r_{\text{failure}}^{\text{buyer}} - r_{\text{success}}^{\text{buyer}}) > 0$$

$$\text{つまり, } p_2^{\text{seller}} > 0 \text{ かつ } r_{\text{failure}}^{\text{buyer}} - r_{\text{success}}^{\text{buyer}} > 0$$

$$p_2^{\text{seller}} > 0 \quad (5.1)$$

$$0 > r_{\text{success}}^{\text{buyer}} - r_{\text{failure}}^{\text{buyer}} \quad (5.2)$$

$$E(R|s_1^{\text{buyer}}) > E(R|s_3^{\text{buyer}})$$

$$\therefore p_1^{\text{seller}} (\text{goods} + r_{\text{success}}^{\text{buyer}}) + p_2^{\text{seller}} r_{\text{failure}}^{\text{buyer}} > p_1^{\text{seller}} (\text{goods} + r_{\text{failure}}^{\text{buyer}}) + p_2^{\text{seller}} r_{\text{success}}^{\text{buyer}}$$

$$\therefore p_1^{\text{seller}} (r_{\text{success}}^{\text{buyer}} - r_{\text{failure}}^{\text{buyer}}) > 0$$

$$\text{つまり, } p_1^{\text{seller}} > 0 \text{ かつ } r_{\text{success}}^{\text{buyer}} - r_{\text{failure}}^{\text{buyer}} > 0$$

$$p_1^{\text{seller}} > 0 \quad (5.3)$$

$$r_{\text{success}}^{\text{buyer}} - r_{\text{failure}}^{\text{buyer}} > 0 \quad (5.4)$$

$$E(R|s_1^{\text{buyer}}) > E(R|s_4^{\text{buyer}})$$

$$\therefore p_1^{\text{seller}} (\text{goods} + r_{\text{success}}^{\text{buyer}}) + p_2^{\text{seller}} r_{\text{failure}}^{\text{buyer}} > p_1^{\text{seller}} (\text{goods} + r_{\text{failure}}^{\text{buyer}}) + p_2^{\text{seller}} r_{\text{success}}^{\text{buyer}}$$

$$\therefore p_1^{\text{seller}} (r_{\text{success}}^{\text{buyer}} - r_{\text{failure}}^{\text{buyer}}) + p_2^{\text{seller}} (r_{\text{failure}}^{\text{buyer}} - r_{\text{success}}^{\text{buyer}}) > 0$$

$$(p_1^{\text{seller}} - p_2^{\text{seller}}) (r_{\text{success}}^{\text{buyer}} - r_{\text{failure}}^{\text{buyer}}) > 0$$

$$(p_1^{\text{seller}} - p_2^{\text{seller}}) (r_{\text{success}}^{\text{buyer}} - r_{\text{failure}}^{\text{buyer}}) > 0 \quad (5.5)$$

の 3 つを満たす必要がある。しかし、(5.2) と (5.4) は矛盾するため、 $(r_{\text{success}}^{\text{seller}}, r_{\text{failure}}^{\text{seller}}, r_{\text{success}}^{\text{buyer}}, r_{\text{failure}}^{\text{buyer}})$  がいかなる実数の組でも条件①は成り立たない。

### 条件②が成り立たないことの証明

また、条件②の必要条件である  $E(R|s_1^{\text{seller}}) > E(R|s_2^{\text{seller}})$  かつ  $E(R|s_2^{\text{buyer}}) > E(R|s_1^{\text{buyer}})$  かつ  $E(R|s_2^{\text{buyer}}) > E(R|s_3^{\text{buyer}})$  かつ  $E(R|s_2^{\text{seller}}) > E(R|s_4^{\text{buyer}})$  を満たすためには、  
 $E(R|s_2^{\text{buyer}}) > E(R|s_1^{\text{buyer}})$

$$\therefore p_1^{\text{seller}} (\text{goods} + r_{\text{success}}^{\text{buyer}}) + p_2^{\text{seller}} r_{\text{success}}^{\text{buyer}} > p_1^{\text{seller}} (\text{goods} + r_{\text{success}}^{\text{buyer}}) + p_2^{\text{seller}} r_{\text{failure}}^{\text{buyer}}$$

$$\therefore p_2^{\text{seller}}(r_{\text{success}}^{\text{buyer}} - r_{\text{failure}}^{\text{buyer}}) > 0$$

$$0 \leq p_2^{\text{seller}} \leq 1 \text{ より,}$$

$$p_2^{\text{seller}} > 0 \text{ かつ } r_{\text{success}}^{\text{buyer}} - r_{\text{failure}}^{\text{buyer}} > 0$$

$$p_2^{\text{seller}} > 0 \tag{5.6}$$

$$r_{\text{success}}^{\text{buyer}} - r_{\text{failure}}^{\text{buyer}} > 0 \tag{5.7}$$

$$E(R|s_2^{\text{buyer}}) > E(R|s_3^{\text{buyer}})$$

$$\therefore p_1^{\text{seller}}(\text{goods} + r_{\text{success}}^{\text{buyer}}) + p_2^{\text{seller}}r_{\text{success}}^{\text{buyer}} > p_1^{\text{seller}}(\text{goods} + r_{\text{failure}}^{\text{buyer}}) + p_2^{\text{seller}}r_{\text{failure}}^{\text{buyer}}$$

$$\therefore p_1^{\text{seller}}(r_{\text{success}}^{\text{buyer}} - r_{\text{failure}}^{\text{buyer}}) + p_2^{\text{seller}}(r_{\text{success}}^{\text{buyer}} - r_{\text{failure}}^{\text{buyer}}) > 0$$

$$\therefore (p_1^{\text{seller}} + p_2^{\text{seller}})(r_{\text{success}}^{\text{buyer}} - r_{\text{failure}}^{\text{buyer}}) > 0$$

$$p_1^{\text{seller}} + p_2^{\text{seller}} = 1 > 0 \text{ より,}$$

$$r_{\text{success}}^{\text{buyer}} - r_{\text{failure}}^{\text{buyer}} > 0$$

$$r_{\text{success}}^{\text{buyer}} - r_{\text{failure}}^{\text{buyer}} > 0 \tag{5.8}$$

$$E(R|s_2^{\text{buyer}}) > E(R|s_4^{\text{buyer}})$$

$$\therefore p_1^{\text{seller}}(\text{goods} + r_{\text{success}}^{\text{buyer}}) + p_2^{\text{seller}}r_{\text{success}}^{\text{buyer}} > p_1^{\text{seller}}(\text{goods} + r_{\text{failure}}^{\text{buyer}}) + p_2^{\text{seller}}r_{\text{success}}^{\text{buyer}}$$

$$\therefore p_1^{\text{seller}}(r_{\text{success}}^{\text{buyer}} - r_{\text{failure}}^{\text{buyer}}) > 0$$

$$p_1^{\text{seller}} > 0 \text{ かつ } r_{\text{success}}^{\text{buyer}} - r_{\text{failure}}^{\text{buyer}} > 0$$

$$p_1^{\text{seller}} > 0 \tag{5.9}$$

$$r_{\text{success}}^{\text{buyer}} - r_{\text{failure}}^{\text{buyer}} > 0 \tag{5.10}$$

ここで, (5.7), (5.8), (5.10) を満たす  $(r_{\text{success}}^{\text{buyer}}, r_{\text{failure}}^{\text{buyer}})$  の組はシステムから決定できるため,

$$p_2^{\text{seller}} > 0 \text{ かつ } p_1^{\text{seller}} > 0 \text{ であれば,}$$

$E(R|s_2^{\text{buyer}}) > E(R|s_1^{\text{buyer}})$  かつ  $E(R|s_2^{\text{buyer}}) > E(R|s_3^{\text{buyer}})$  かつ  $E(R|s_2^{\text{buyer}}) > E(R|s_4^{\text{buyer}})$  を満たすことができる。

ここで仮に  $p_2^{seller} > 0$  かつ  $p_1^{seller} > 0$  が成り立ち  $buyer$  が戦略  $s_2^{buyer}$  を選択するととする。

このとき、 $buyer$  が各戦略をとる確率  $(p_1^{buyer}, p_2^{buyer}, p_3^{buyer}, p_4^{buyer})$  は  $(0, 1, 0, 0)$  と表せる。

ここで、 $p_1^{buyer} = p_4^{buyer} = 0$  のため、 $E(R|s_1^{seller}) > E(R|s_2^{seller})$  は成り立たない。

それゆえに、条件②は成り立たない。

以上より、条件①もしくは条件②のいずれかの条件を満たす  $(r_{success}^{seller}, r_{failure}^{seller}, r_{success}^{buyer}, r_{failure}^{buyer})$  の組を「商取引システム」から決定することはできない。

## 5.8 結論

本稿では、商取引において不正行為を防止するためには、「行動観察不可の条件」を満たした「第 3 者に依存しない仲介システム」が存在している必要があることを論じた。その上で、そのシステムが存在するのかを検証するために、下記の条件を付随した「商取引システム」と、それを仲介とした「商取引ゲーム」を定義して、商取引で不正行為を防止できるインセンティブ設計が可能かを確かめた。

- システムは  $buyer$  によって報告された商取引の結果を観察できる。
- システムは  $seller$  と  $buyer$  がどの戦略を選んだかはわからない。
- 商取引に参加する  $player$  は合理的に (利得の期待値が最も高い) 戦略を決定する。
- システム内には追跡可能な通貨が存在しており、商取引にはその通貨が用いられる。
- システムからは  $player$  の通貨の保有量を操作することができる。

その結果として、不正行為が防止されるための 2 つの戦略組のいずれかに  $seller$  と  $buyer$  の戦略を帰着させるような利得の組を商取引システムから決定することはできないことを証明した。これはつまり、人々は合理的であるという仮定の上で成り立つ「商取引ゲーム」において、不正行為を防止することが不可能であることを意味する。

社会契約は複数の商取引の連鎖と捉えることができる。デイビッドヒュームは同じように考えて、商取引において双方にメリットがあるから社会契約は成立すると考えた。囚人のジレンマから必ずしもパレート最適な答えにたどり着くわけではないことがわかる。商取引において、不正を防止するインセンティブ設計は可能か？ 無法地帯での無人売買機問題商取引システム

## 第6章 倫理ある商取引ゲーム

前章では合理的なプレイヤー達による「商取引ゲーム」において、不正を防止するインセンティブ設計が不可能であることを示した。しかしながら、現実に住む我々は商取引で必ず不正に合うわけではなく、ある程度は不正が防止されている。本章では、この理論と現実の差異は限定合理性にあると考え、そうした戦略をとるプレイヤーのみで構成されたときに不正を防止することができる「倫理ある商取引ゲーム」をモデリングする。そして、そのモデルが全ての戦略をとりえるプレイヤーによって構成された「商取引ゲーム」においても、プレイヤーの構成次第で不正を防止するインセンティブ設計として機能することを示す。

### 6.1 本章における問題提起

前章では商取引ゲームにおいて必ず不正が防止できるようなインセンティブの設計が不可能であることを示した。しかしながら、現実的に私達は商取引で不正行為に遭遇することは稀である。なぜ先の理論では商取引で不正を防止するインセンティブ設計ができないにも関わらず、現実の生活の中で我々は商取引を成功させることができているのだろうか。

### 6.2 本章の仮説

この問いの答えは、限定合理性 (Simon 1947[7]) によって戦略の均衡点に変化し、不正を防止できるインセンティブ設計が可能になっているためだと思われる。限定合理性とは、認知能力の限界によって意思決定主体が限られた合理性しか持ち得ないことである。各プレイヤーにとっての最適な戦略は他のプレイヤーがとる戦略に依存しているため、当然、一部のプレイヤーが合理的に戦略を選ばないのであれば最適な戦略の均衡点も変化する。それによって、全てのプレイヤーが合理的である「商取引ゲーム」においては不可能だった相手が不正を行う可能性を推定することが可能になり、不正を防止するインセンティブ設計が可能になると予想される。

### 6.3 提案手法

このような限定合理性を「倫理」と呼び、不正行為にあった場合に必ず「失敗」を報告する行動規範とする。また、この倫理に従うプレイヤーのみによって構成される「商取引ゲーム」を「倫理ある商取引ゲーム」とする。本章では、「倫理ある商取引ゲーム」において、不正が防止されるインセンティブ設計を行い、そのインセンティブ設計を通常の「商取引ゲーム」にも適用することで、プレイヤーの構成によっては不正を防止することが可能になることを示す。

#### 6.3.1 倫理ある商取引ゲーム

「倫理ある商取引ゲーム」とは、*seller* が不正を行った場合に *buyer* がかならず「失敗」を報告する「商取引ゲーム」である。このゲームのゲーム木と非協力戦略型ゲームの利得は図 6.2 と表 6.1 のように表せる。

### 6.3.2 誠実な戦略をとった割合と成功が報告される割合の関係

「倫理ある商取引ゲーム」においては、全てのプレイヤーは「倫理」に従っているため、商取引の真の成功率は「商取引システム」に「成功」が報告された割合以上となる。つまり、任意のプレイヤー  $p$  と  $q$  が過去に「商取引ゲーム」を行った際に、誠実な戦略 ( $s_1^{seller}$  もしくは  $s_1^{buyer}$ ) をとってきた割合  $HonestStrategyRate(p, q)$  と、成功が報告された割合  $ReportedSuccessRate(p, q)$  について、次の関係がいえる。

$$HonestStrategyRate(p, q) \geq ReportedSuccessRate(p, q) \quad (6.1)$$

## 6.4 「倫理ある商取引ゲーム」において不正が防止される条件

### 6.4.1 不正が抑制される戦略組と期待利得の不等式

「倫理ある商取引ゲーム」において、不正を防止するためには、*seller* と *buyer* の戦略組を ( $s_1^{seller}, s_1^{buyer}$ ) に帰着させる必要がある。そのためには *seller* が戦略  $s_1^{seller}$  をとった場合の期待利得  $E(r|s_1^{seller})$  が戦略  $s_2^{seller}$  をとった場合の期待利得  $E(r|s_2^{seller})$  より大きく、*buyer* が戦略  $s_1^{buyer}$  をとった場合の期待利得  $E(r|s_1^{buyer})$  が戦略  $s_3^{buyer}$  をとった場合の期待利得  $E(r|s_3^{buyer})$  より大きくならなければならない。つまり、 $E(r|s_1^{seller}) > E(r|s_2^{seller})$  かつ  $E(r|s_1^{buyer}) > E(r|s_3^{buyer})$  を満たす ( $r_{success}^{seller}, r_{failure}^{seller}, r_{success}^{buyer}, r_{failure}^{buyer}$ ) の組を「商取引システム」から決定できる必要がある。

### 6.4.2 *seller* と *buyer* の期待利得

*seller* と *buyer* の各戦略の利得の期待値は以下のように表せる。

$$E(R|s_1^{seller}) = p_1^{buyer}(r_{success}^{seller} + \epsilon^{seller}) + p_3^{buyer}(r_{failure}^{seller} + \lambda^{seller})$$

$$E(R|s_2^{seller}) = p_1^{buyer}(goods + r_{failure}^{seller} + \lambda^{seller}) + p_3^{buyer}(goods + r_{failure}^{seller} + \lambda^{seller})$$

$$= goods + r_{failure}^{seller} + \lambda^{seller}$$

$$\because p_1^{buyer} + p_3^{buyer} = 1 \text{ (「倫理ある商取引ゲーム」において、} p_2^{buyer} \text{ と } p_4^{buyer} \text{ は 0 であるため)}$$

$$E(R|s_1^{buyer}) = p_1^{seller}(goods + r_{success}^{buyer} + \epsilon^{buyer}) + p_2^{seller}(r_{failure}^{buyer} + \lambda^{buyer})$$

$$E(R|s_3^{buyer}) = p_1^{seller}(goods + r_{failure}^{buyer} + \lambda^{buyer}) + p_2^{seller}(r_{failure}^{buyer} + \lambda^{buyer})$$

$$= p_1^{seller} goods + r_{failure}^{buyer} + \lambda^{buyer}$$

$$\because p_1^{seller} + p_2^{seller} = 1$$

### 6.4.3 *seller* が誠実な戦略をとる条件

$$E(R|s_1^{seller}) > E(R|s_2^{seller})$$

$$\therefore p_1^{buyer}(r_{success}^{seller} + \epsilon) + p_2^{buyer}(r_{failure}^{seller} + \lambda) > goods + r_{failure}^{seller} + \lambda$$

$$\therefore p_1^{buyer}(r_{success}^{seller} + \epsilon) - p_1^{buyer}(r_{failure}^{seller} + \lambda) > goods$$

$$\therefore p_1^{buyer}(r_{success}^{seller} - r_{failure}^{seller} + \epsilon - \lambda) > goods$$

仮定より,  $\epsilon > \lambda$  のため,  $p_1^{buyer}(r_{success}^{seller} - r_{failure}^{seller}) \geq goods$  を満たせばよい.

$0 < p_1^{buyer}$  を仮定するならば,  $r_{success}^{seller} - r_{failure}^{seller} \geq \frac{goods}{p_1^{buyer}}$

### 6.4.4 *buyer* が誠実な戦略をとる条件

$$E(R|s_1^{buyer}) > E(R|s_3^{buyer})$$

$$\therefore p_1^{seller}(goods + r_{success}^{buyer}) + p_2^{seller}r_{failure}^{buyer} > p_1^{seller}(goods + r_{failure}^{buyer}) + p_2^{seller}r_{failure}^{buyer}$$

$$\therefore p_1^{seller}(r_{success}^{buyer} - r_{failure}^{buyer}) > 0$$

$0 < p_1^{seller}$  を仮定するならば,

$$r_{success}^{buyer} - r_{failure}^{buyer} > 0$$

上記をまとめると,  $0 < p_1^{buyer}$  かつ  $0 < p_1^{seller}$  を仮定した上で,

$$r_{success}^{seller} - r_{failure}^{seller} \geq \frac{goods}{p_1^{buyer}} \text{ かつ } r_{success}^{buyer} - r_{failure}^{buyer} > 0$$

を満たせば, 「倫理ある商取引ゲーム」で不正を防止することができる.

### 6.4.5 信頼度 $p^{player}$

ここで, 任意の *player* が誠実な戦略 ( $p_1^{seller}, p_1^{buyer}$  のいずれか) をとる主観確率を  $p_1^{player}$  とすると,  $p_1^{player}$  は各プレイヤーに対して誠実な戦略をとった割合  $HonestyStrategyRate(player, opportunity)$  と任意の重み  $w^{player}$  を用いて次のように表せる.

$$p_1^{player} \equiv \sum_{opp}^{players} w^{opp} HonestyStrategyRate(player, opp) \quad (6.2)$$

#### 6.4.6 最低信頼度 $T^{player}$

しかし、「商取引システム」からは  $HonestyStrategyRate(player, opportunity)$  は未知のため、信頼度  $p_1^{player}$  を求めることができない。そこで  $HonestyStrategyRate$  の代わりに  $ReportedSuccessRate$  を用い、信頼度  $p^{player}$  を計算するのと同じ重み  $w^{player}$  の荷重総和をとったものを、最低信頼度  $T^{player}$  と定義する。

$$T^{player} \equiv \sum_{opp}^{players} w^{opp} ReportedSuccessRate(player, opp) \quad (6.3)$$

#### 6.4.7 最低信頼度を用いた条件

ここで  $HonestyStrategyRate \geq ReportedSuccessRate$  であるため、 $p_1^{player} \geq T^{player}$  がいえる。ゆえに、

$$r_{success}^{seller} - r_{failure}^{seller} \geq \frac{goods}{T^{buyer}} \geq \frac{goods}{p_1^{buyer}} \quad (6.4)$$

となる。

つまり、 $0 < p_1^{buyer}$  かつ  $0 < p_1^{seller}$  を仮定した上で、

$$r_{success}^{seller} - r_{failure}^{seller} \geq \frac{goods}{T^{buyer}} \text{ かつ } r_{success}^{buyer} - r_{failure}^{buyer} > 0 \quad (6.5)$$

を満たす  $(r_{success}^{seller}, r_{failure}^{seller}, r_{success}^{buyer}, r_{failure}^{buyer})$  の組を「商取引システム」から決定できれば、「倫理ある商取引ゲーム」において不正を防止することができる。

### 6.5 実験方法

先の不正が防止される条件を満たすインセンティブ設計を行う「商取引システム」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 8000 回づつ繰り返し、エージェントの構成と step13 で求める「報告された成功率」と「真の成功率」を記録する。

#### 6.5.1 エージェントの種類

下記の 8 タイプのエージェントを用意する。

Type A seller のとき商品を送り、buyer のときは商品を受け取った場合は「成功」、受け取らなかった場合は「失敗」を報告する。

Type B seller のとき商品を送り、buyer のときは商品を受け取った場合は「成功」、受け取らなかった場合は「成功」を報告する。



Type C seller のとき商品を送り、buyer のときは商品を受け取った場合は「失敗」、受け取らなかった場合は「成功」を報告する。

Type D seller のとき商品を送り、buyer のときは商品を受け取った場合は「失敗」、受け取らなかった場合は「失敗」を報告する。

Type E seller のとき商品を送らず、buyer のときは商品を受け取った場合は「成功」、受け取らなかった場合は「失敗」を報告する。

Type F seller のとき商品を送らず、buyer のときは商品を受け取った場合は「成功」、受け取らなかった場合は「成功」を報告する。

Type G seller のとき商品を送らず、buyer のときは商品を受け取った場合は「失敗」、受け取らなかった場合は「成功」を報告する。

Type H seller のとき商品を送らず、buyer のときは商品を受け取った場合は「失敗」、受け取らなかった場合は「失敗」を報告する。

### 6.5.2 試行

step 1 時刻  $t$  を 0 とする。

step 2 「商取引システム」の各エージェントの通貨保有量を 8 とする。

step 3 各プレイヤーが互いに seller と buyer のそれぞれの役割で 1 度ずつ商取引ゲームを行う順序を決定する。(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、seller と buyer を決定する。

step 6 「商取引システム」はこれまで報告された結果から、「成功」と「失敗」が報告された場合の seller と buyer の通貨保有量を計算する。

step 7 seller は自身の戦略に基づいて、buyer に商品を送るか、もしくは商品を送らない。

step 8 buyer は自身の戦略と step 6 の seller の行動に基づいて、「商取引ゲーム」の結果を決定する。

step 9 buyer は決定した結果を「商取引システム」に報告する。

step 10 「商取引システム」は step 6 で計算した通貨保有量がいずれの場合にも 0 未満にならない場合、buyer から報告された結果を記録する。

step 11 step 6 で計算した通貨保有量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、step 4 に戻る。

step 13 過去 56 回の商取引ゲームにおいて、step 10 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」を求める。

## 6.6 「商取引システム」の詳細

本節ではシミュレーションを実装するにあたって必要となる商取引システムの仕様の詳細の一部を紹介する。完全な実装については、GitHub のソースコードを参照。

### 6.6.1 ReputationWeight

最低信頼度  $T^{player}$  を求めるためには、 $ReportedSuccessRate(p, q)$  に係る任意の重み  $w^{player}$  を決定する必要がある。この重み  $w^{player}$  は、任意の  $player$  が誠実な戦略をであろう主観確率を考える際にその  $player$  と取引相手  $opportunity$  との間での報告された成功率  $ReportedSuccessRate(player, opportunity)$  をどの程度信頼するかを表している。本論では、保有している通貨の全体に占める割合を ReputationWeight  $w^{player}$  任意の  $player$  の通貨保有量を  $b^{player}$  としたとき、

$$w^{player} \equiv \frac{b^{player}}{\sum_i^{players} b^i}$$

### 6.6.2 「成功」が報告された場合の通貨保有量の変化

「商取引ゲーム」において、 $buyer$  から成功が報告された場合、 $buyer$  は商品  $goods$  の価格  $price$  だけ通貨保有量が減り、 $seller$  は  $price$  だけ通貨保有量が増えるものとする。

そのため商取引前後では  $seller$  と  $buyer$  の残高の合計は変化しない。ここから、 $r_{success}^{seller}$  と  $r_{success}^{buyer}$  は以下のように記せる。

$$\begin{aligned} r_{success}^{seller} &= price \\ r_{success}^{buyer} &= -price \\ r_{success}^{seller} + r_{success}^{buyer} &= 0 \end{aligned}$$

### 6.6.3 EscrowCost

まずは「失敗」が報告された時に  $seller$  と  $buyer$  から失われる通貨の量の合計を  $EscrowCost$  とおいて考え、同時に商品価格  $price$  にエスクロー係数  $E$  を掛けたものとする。（ここで  $price$  は  $goods$  の価格である）

$$\begin{aligned} EscrowCost &\equiv (r_{success}^{buyer} - r_{failure}^{buyer}) + (r_{success}^{seller} - r_{failure}^{seller}) \\ &= E \cdot price \end{aligned} \tag{6.6}$$

### 6.6.4 EscrowCost の負担比率

$EscrowCost$  の負担比率は  $seller$  と  $buyer$  の最低信頼度  $T^{player}$  を用いる。

$$(r_{success}^{buyer} - r_{failure}^{buyer}) : (r_{success}^{seller} - r_{failure}^{seller}) = T^{buyer} : T^{seller} \tag{6.7}$$

### 6.6.5 EscrowCost の分配

「失敗」が報告されたときに *EscrowCost* が消失すると、全体の通貨量が減少して通貨の価値が上がり商品価格が下がる。商品価格の変動を防ぐために本実験では *seller* と *buyer* 以外の全てのプレイヤーに、そのプレイヤーの通貨保有量に応じて *EscrowCost* を分配する。*seller* と *buyer* を含まないのは、分配によって「商取引ゲーム」のインセンティブ設計が変化しないようにするためである。

## 6.7 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 6.1 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。

## 6.8 結論

実験とその評価を踏まえて、「倫理」という限定合理性を仮定した上でインセンティブ設計を行うことで、プレイヤーの構成によっては「商取引ゲーム」において不正を防止することが可能であることがわかる。また、プレイヤーの構成と不正防止の成功成功の関係性については、先の実験の図のとおりである。

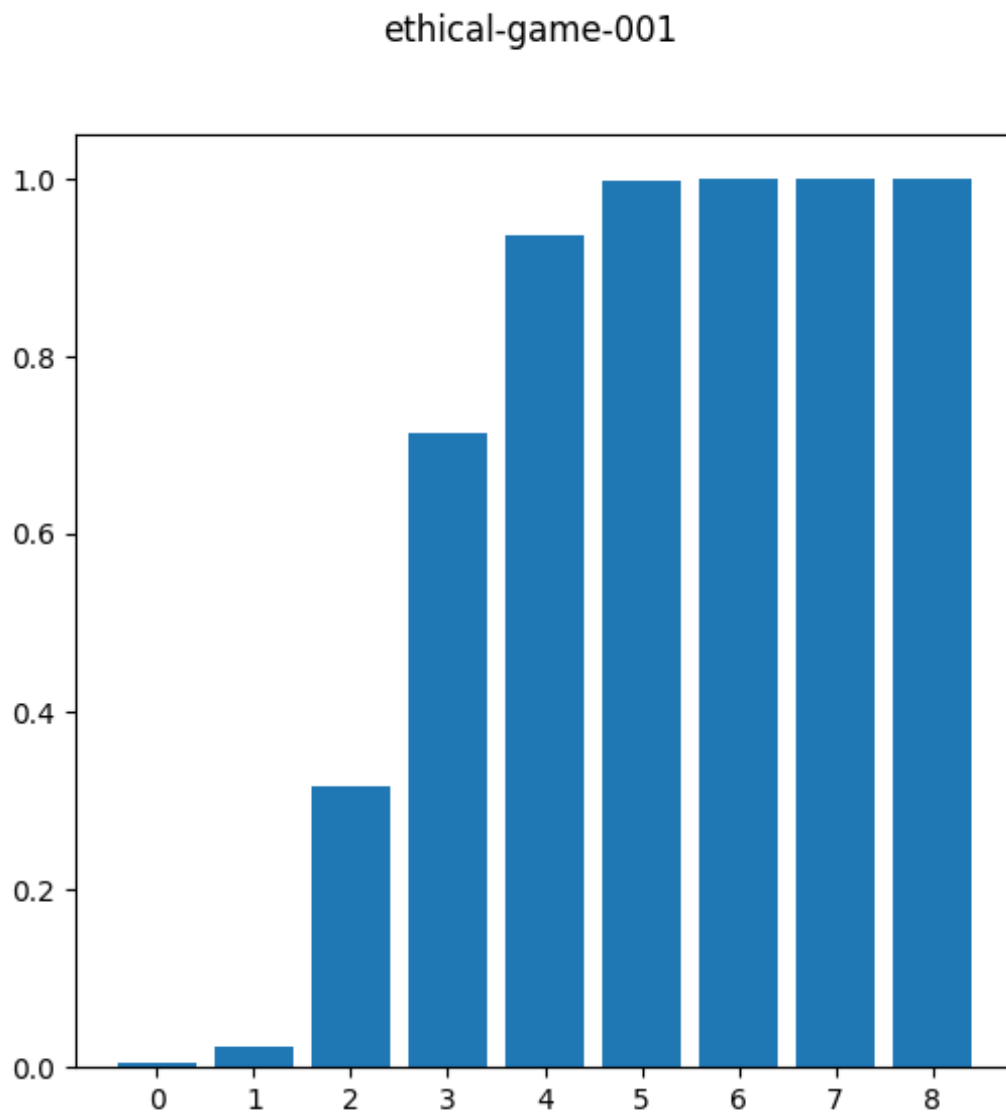


図 6.1: 誠実なエージェントの数と「不正防止の成功」に至った割合

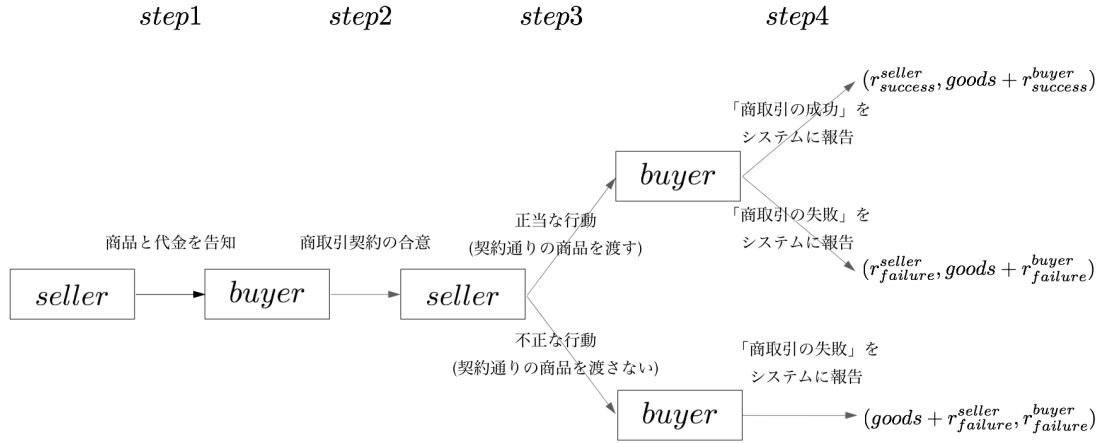


図 6.2: 「倫理ある商取引ゲーム」のゲーム木

		<i>buyer</i>	
		$s_1^{buyer}$	$s_3^{buyer}$
<i>seller</i>	$s_1^{seller}$	$(r_{success}^{seller}, goods + r_{success}^{buyer})$	$(r_{failure}^{seller}, goods + r_{failure}^{buyer})$
	$s_2^{seller}$	$(goods + r_{failure}^{seller}, r_{failure}^{buyer})$	$(goods + r_{failure}^{seller}, r_{failure}^{buyer})$

表 6.1: 非協力戦略型ゲームとして表した「倫理ある商取引ゲーム」の利得表

## 第7章 複雑系としての社会契約

「倫理ある商取引ゲーム」において、倫理という限定合理性を仮定した上で商取引システムのインセンティブ設計を行えば、プレイヤーの構成によっては不正が防止され、商取引契約の内容が果たされるようになることがわかった。本章では、この商取引契約の内容に焦点を当て、これまで外部の強制執行力として存在が仮定されていた「商取引システム」を、ある集団の成員が構成する複雑系の内部で自己組織化されたシステムとして再現する商取引契約の内容について提案し、マルチエージェントシミュレーションによって、商取引システムが存在しない場合でも、社会契約が成立しうることを示す。

### 7.1 本章で取り組む問題

本章では先の章で扱った「倫理ある商取引ゲーム」において、外部の強制執行力として存在が仮定されていた「商取引システム」が存在しない場合でも、プレイヤーの構成によって「商取引ゲーム」の不正を防止することが可能であるかという問いに取り組む。

### 7.2 本章で検証する仮説

商取引システムが存在する場合に、商取引契約を履行させることが可能であるとする。このとき、各プレイヤーが商取引システムと同様の役割を演じその正当性を保証する商取引契約を結ぶことができれば、外部の強制執行力としての商取引システムが存在せずとも相互監視の元で商取引ゲームで不正を防止し続けることが可能となるだろう。

### 7.3 提案手法

先の仮説を検証するためには、各プレイヤーが商取引システムの役割を正しく演じる商取引契約の内容を記述し、その商取引契約の履行を約束する「商取引ゲーム」を繰り返した結果、全てのプレイヤーが「商取引ゲーム」で不正を行えない状態になりうることを示す必要がある。

ここで、「商取引システム」とはどういったシステムであったかに立ち返ると、「商取引システム」とは、各プレイヤーの初期の通貨保有量と保存された各時刻の商取引の記録から、各プレイヤーの通貨保有量を一意に決定するシステムである。それ故、初期の通貨保有量と保存された各時刻の商取引の記録が全てのプレイヤーで一致しているとき、全てのプレイヤーが商取引システムが正しく動作しているといえる。

それを実現するためには、各プレイヤーが保存している過去の商取引の記録を互いに確認し合い、差異が生じた場合に全てのプレイヤーに「失敗」を報告する商取引契約の内容を記述すればよい。

そこで、本章では下記のような『「商取引システム」を自己組織化する商取引契約』を提案し、先の仮説が成り立つかを検証する。

### 7.3.1 「商取引システム」を自己組織化する商取引契約

step 1 時刻  $t$  を 0 とする。

step 2 プレイヤーの人数と各プレイヤーが初期に保有する通貨の量を任意に決定する。

step 3  $N$  人のプレイヤーが互いに全てのプレイヤーと seller と buyer の 2 つの役割で商取引ゲームを行う周期ある順序を決定する。この周期の長さは  $N \cdot N - 1$  となる。

step 4 時刻  $t$  を 1 進める。

step 5 step3 で決定した順序に基づいて時刻  $t$  の seller と buyer を決定する。

step 6 自分が seller ならば、時刻  $\max\{0, t - n * (n - 1)\}$  から時刻  $t$  までに報告された商取引ゲームの記録のうち、報告者がその商取引ゲームの記録の buyer と一致するものを全て buyer に送信する。

step 7 自分が buyer ならば、seller から受け取った商取引ゲームの記録をメモする。この際、記録する商取引の記録の報告者を seller に書き換える。

step 8 自分が buyer ならば、時刻  $\max\{0, t - 2 * n * (n - 1)\}$  から時刻  $\max\{0, t - n * (n - 1)\}$  までの各時刻について、メモされた商取引の記録のうち自身が報告者の記録はそのまま商取引システムに保存する。それ以外の記録は、「成功」と「失敗」のそれぞれの報告者について、自身の所有する商取引システムの通貨保有量の総和をとり、大きい方の結果を採用して商取引システムに保存する。

step 9 自身が buyer ならば、支持する商取引の記録を自身の所有する商取引システムに保存する。

step 10 自身が buyer ならば、時刻  $\max\{0, t - 2 * n * (n - 1)\}$  から時刻  $\max\{0, t - n * (n - 1)\}$  までの各時刻について、自身の所有する商取引システムに保存した商取引の記録と、報告者が seller であるメモされた商取引の記録を比較し、全てが一致する場合には「成功」を 1 つでも一致しない場合は「失敗」を全てのプレイヤーに報告する。

step 11 全てのプレイヤーは buyer から受けた商取引の記録をメモする。この際、商取引の記録の報告者を buyer に書き換える。

step 12 step4 に戻る。

## 7.4 実験方法

明確に商取引契約を履行しているプレイヤーとそうでないプレイヤーを分けるために、『「商取引システム」を自己組織化する商取引契約』の内容を一部改変した「実験用の商取引契約」と 8 タイプのエージェントを用意する。

誠実なエージェント (エージェントタイプ A) の人数  $K$  が 0~7 までの場合でそれぞれ 8000 回づつ下記の施行を繰り返す。

step 1 タイプ A のエージェントの人数が  $K$  になり、残りが B H で構成されるような 8 体のエージェントをランダムに用意する。

step 2 初期の各プレイヤーの通貨保有量を 8 とし、時刻  $t$  が 1120 になるまで実験用の商取引契約に基づく商取引ゲームを繰り返す。

step 3 時刻  $t = 1120$  まで下記の施行を繰り返し、各エージェントの所有する商取引システムが有効だと判断したについて、報告された成功率と真の成功率を記録する。

### 7.4.1 実験用の商取引契約

**step 6** 自分が *seller* ならば、時刻  $\max\{0, t - n * (n - 1)\}$  から時刻  $t$  までに報告された商取引ゲームの記録のうち、報告者がその商取引ゲームの記録の *buyer* と一致するものを全て *buyer* に送信する。その際、メッセージに「UTC」の文字列を追記して送信する。

**step 10** 自身が *buyer* ならば、時刻  $\max\{0, t - 2 * n * (n - 1)\}$  から時刻  $\max\{0, t - n * (n - 1)\}$  までの各時刻について、自身の所有する商取引システムに保存した商取引の記録と、報告者が *seller* であるメモされた商取引の記録を比較し、全てが一致する場合には「成功」を1つでも一致しない場合は「失敗」を全てのプレイヤーに報告する。ただし、step7で *seller* から送られてきたメッセージが「UTC」でない場合は、全てが一致していても「失敗」を報告する。

### 7.4.2 8種類のエージェント

- タイプ A…step6で「UTC」を送り、step10で「UTC」が送られてきた場合は「成功」、それ以外の場合は「失敗」を報告する
- タイプ B…step6で「UTC」を送り、step10で「UTC」が送られてきた場合は「成功」、それ以外の場合は「成功」を報告する
- タイプ C…step6で「UTC」を送り、step10で「UTC」が送られてきた場合は「失敗」、それ以外の場合は「成功」を報告する
- タイプ D…step6で「UTC」を送り、step10で「UTC」が送られてきた場合は「失敗」、それ以外の場合は「失敗」を報告する
- タイプ E…step6で「JST」を送り、step10で「JST」が送られてきた場合は「成功」、それ以外の場合は「失敗」を報告する
- タイプ F…step6で「JST」を送り、step10で「JST」が送られてきた場合は「成功」、それ以外の場合は「成功」を報告する
- タイプ G…step6で「JST」を送り、step10で「JST」が送られてきた場合は「失敗」、それ以外の場合は「成功」を報告する
- タイプ H…step6で「JST」を送り、step10で「JST」が送られてきた場合は「失敗」、それ以外の場合は「失敗」を報告する

## 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。



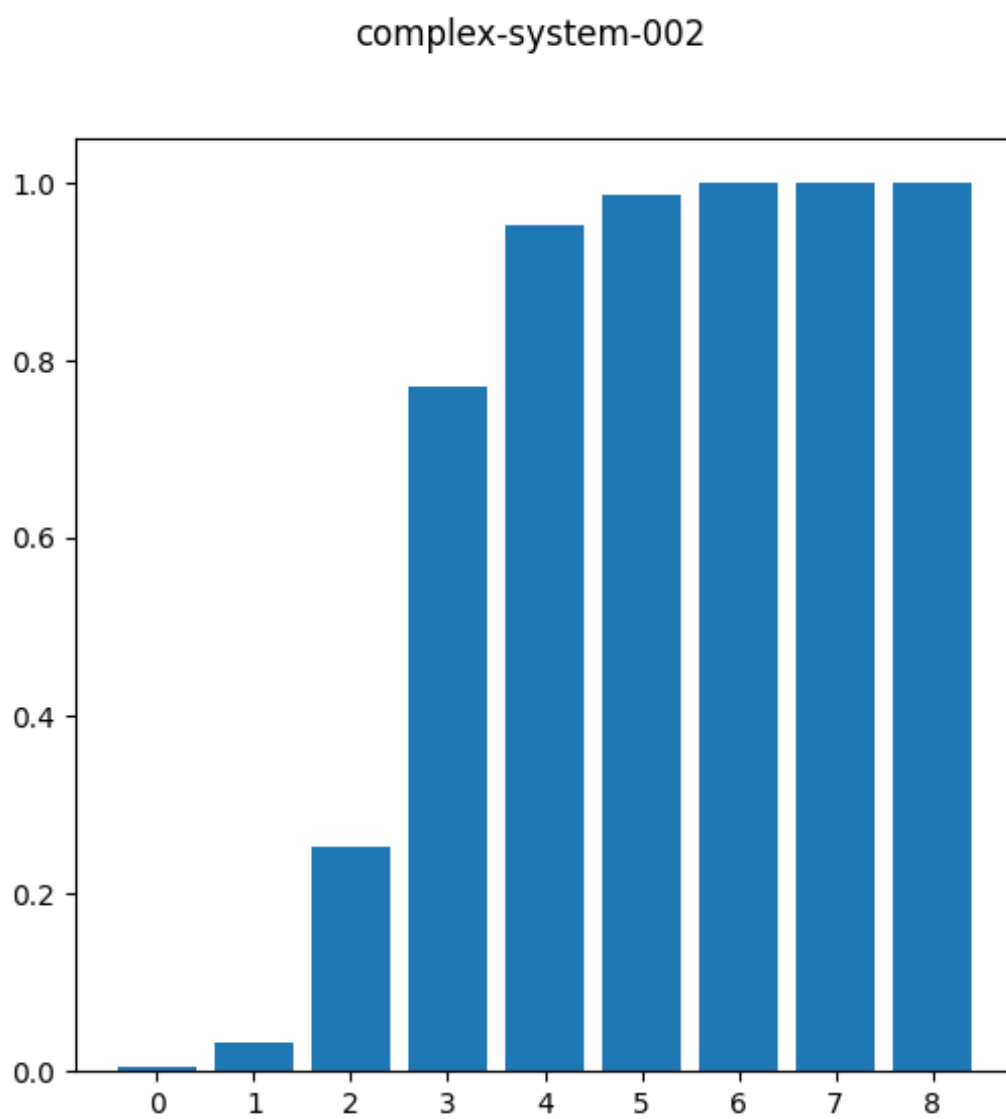


図 7.1: 誠実なエージェントの数と自己組織化に成功した割合

## 第8章 結論

### 8.1 本論のまとめ

### 8.2 本論の課題

### 8.3 今後の研究

## 謝辭

## 参考文献

- [1] 社会契約について, デイビットヒューム,
- [2] 人類不平等起源説
- [3] Binmore, Ken. 1994. Playing Fair (Game Theory and the Social Contract; vol 1). Cambridge: The MIT Press.
- [4] Binmore, Ken. 1998. Just Playing (Game Theory and the Social Contract, vol. 2). Cambridge: The MIT Press.
- [5] Binmore, Ken. 2005. Natural Justice. New York: Oxford University Press.
- [6] Brian Skyrms, The Stag Hunt, 2001.
- [7] Simon, H. A., "Administrative Behavior", 1947.
- [8] Leslie Lamport, Robert Shostak, Marshall Pease: "The Byzantine Generals Problem", ACM Transactions on Programming Languages and Systems, Vol.4, No.3, pp. 382-401, 1982.
- [9] 倫理ある商取引ゲームの実験コード      <https://github.com/nontangent/social-contract/tree/master/experiments/ethical-game/01>