

# 安心装置のもとでの協力の発生： 進化シミュレーションによる分析

高木英至

埼玉大学 教養学部

NCC00521@nifty.ne.jp

キーワード：安心、信頼、協力、社会的ディレンマ、進化シミュレーション

【要約】 $n$ 人協力関係の発生経路には2種類が考えられる。第1は「水平的」経路である。エージェントが相手の協力可能性を見きわめながら協力確率の高い相手だけを関係に含める戦略（選択的受容戦略）をとる場合が該当する。第2は「垂直的」経路である。科罰能力を持つ安心請負人と契約し、安心請負人が非協力を抑制することを見込んで協力関係を形成する場合である。前者は「信頼」に基づく協力、後者は「安心」による協力と呼ぶことができる。筆者の以前のシミュレーションは、ある程度の裏切り誘因があるときに高い信頼とともに選択的受容戦略が進化し、信頼に基づく協力が出現できることを示した。本研究では、安心請負人の出現を許す状況では、裏切り誘因が強く「水平的」経路では協力が出現できない条件でも、安心に基づく協力が出現できることを示す。

## 1 はじめに：信頼に基づく協力

本研究の背後にある問題関心は、社会の基本的な構造もしくは秩序がいかに創発するか、という点にある。社会科学は従来、社会がある程度組織化された基本構造を持つことを自明の所与と前提し、その基本構造がなぜ出現しているかを議論の対象とはしなかった。筆者がする意図するのは、社会の基本構造の出現の論理を見出すことにより、基本構造が持つ社会への含意を解明することである。

社会の基本構造の1つはその中に一定範囲での協力が存在し、その協力の「基盤」が備わっていることである。しかし協力とは潜在的に社会的ディレンマの状況にある。それゆえ、協力が出現することは自明とはいえない。

これまで多くの場合、協力の発生は2人の囚人のデ

ィレンマ（2PD）によりモデル化されてきた。2PDに基づく協力が社会の重要な構成要素であるのは確かである。しかし、社会的により意味のある協力は $n$ 人関係として生じやすいと考えるべきだろう。また、2PDは、2者間という限定された範囲で相手の利得を直接コントロールできる場合である。つまり当事者間の直接的な相互作用によって協力を形成しやすいという意味で特殊な状況にある。

本研究が取り組むのは $n$ 人協力関係の発生である。 $n$ 人協力関係は内在するディレンマが強いため、2人関係の場合とは異なった考慮を要する可能性がある。

$n$ 人協力関係の発生について、筆者は以前に「協力呼びかけゲーム」に基づくシミュレーションを実施している（[3]）。

協力呼びかけゲームでは、社会を構成するエージェントが順番に協力の呼びかけ人になる機会が与えられる。呼びかけ人はその戦略に応じて、ランダムに相手を選んで呼びかけることもできるし（ランダム申込み）、相手の過去の協力履歴（協力率）から相手を選んで呼びかけることもできる（選択的申込み）。呼びかけられたエージェントも呼びかけにランダムな確率で応じることもできるし（ランダム受諾）、呼びかけ人と呼びかけられた他者の協力履歴を考慮して応じるか否かを判断することもできる（選択的受諾）。申込みと受諾が選択的な戦略が「選択的受容戦略」である。選択的な申込み／受諾をするとき、過去の協力率が得られない相手については「信頼」、つまり未知の相手の協力確率の推定値を代入して判断が生じると考える。なお、協力関係は社会的ディレンマ状況であり（[2]）、裏切りは利得をもたらすとともに他の参加者の利得を低下させる。

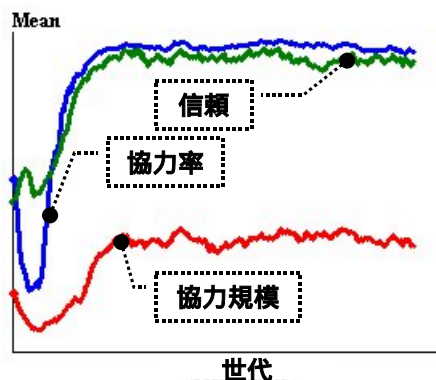


図1:協力の進化(最初の250世代)

協力呼びかけゲームに基づくシミュレーションでは協力の進化を観測できた。世代進行とともに選択的受容戦略が優越し、協力率、協力関係の規模、信頼が上昇したのである(図1)。

このシミュレーションで裏切り誘因を操作したときの結果を示すのが図2である。図2では、裏切り誘因が最大の条件で協力が崩れるけれども、他の条件では裏切り誘因にかかわらず協力率と協力関係規模が維持されているのが分かる。

興味深いのは信頼の値の挙動である。協力が崩壊する裏切り誘因最大の条件を除いて、協力が維持されている範囲では、裏切り誘因が高い方が信頼が高い。この結果は次のように理解できる。ある程度の裏切りがある状況ではエージェントは選択的受容戦略によって協力相手を選別することで利得を高めることができる。ここで問題は、協力確率が分からない相手に出会う可能性が高いことである。未知の相手の協力確率には信頼が代入されるとすれば、信頼が低いと協力関係を形成しにくく、したがって利得を高める機会を失うことになる。そのため高信頼と選択的受容戦略が結びつき、裏切り誘因の高い条件で信頼が上がることになる。他方、裏切り誘因の低い状況は自然に協力が生じる、いわば「安心」状況にある。この安心状況ではエージェントは選択的受容戦略を選択する理由もなく、したがって信頼も低いままとなると考えられる。

## 2 安心装置問題

以上の結果が出た時点で筆者が関心を抱いたのは、協力が崩壊した条件、つまり裏切り誘因がきわめて高い条件でいかにして協力を生じさせるかであった。この条件では裏切り誘因が強くても、なおも社会的ディ

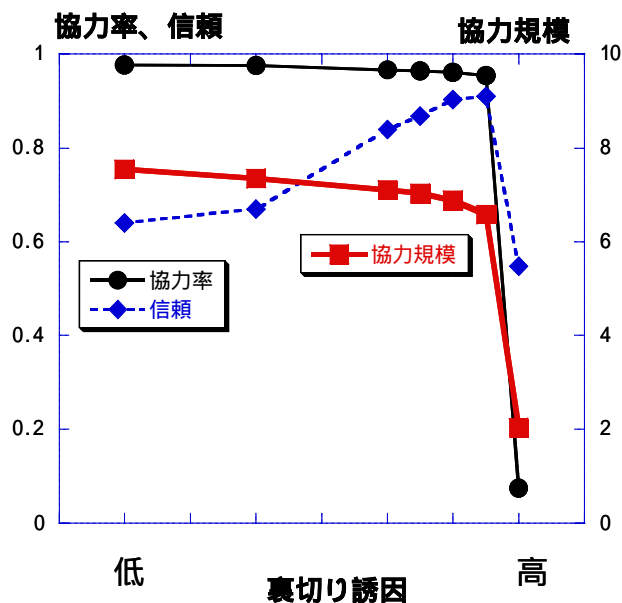


図2:裏切り誘因の効果

レンマの利得構造が保持されている。つまり相互協力は潜在的には参加者の利益につながる。しかしここまでの前提では、裏切り誘因最大の条件で協力を高める方法は見当たらない。例えばエージェントの信頼を高めたり、specific trust(特定の他者の協力確率推定)を高めたりしても効果はなかった。

1つの考慮は、ここまでの議論では協力発生「水平的経路」しか考えて来なかったことである。つまり、同等なエージェント間の相互作用の中から生まれる戦略に依存した協力発生「垂直的経路」もあるだろう。権力者に相当する「安心請負人」が登場し、裏切り者を罰することで協力を保証するような「安心装置」ができ上がることである。

ではこのような安心装置はいかにして bottom-up に出現できるのか？

まず考えられるのは、エージェントが「裏切ったら罰する」ような「制裁同盟」を結び、裏切り者に罰を科すことである。しかし安心装置としての制裁同盟は、制裁自体への free-riding が生じるため、Axelrod[1]の Norm Game が示すように、崩壊する運命にある。

安心装置としての制裁同盟の弱点は、その負担が社会的ディレンマを構成する点にある。したがって誰かを「安心請負人」にし、その安心請負人との取引として設立するなら安心装置は維持できるかも知れない。安心請負人は「裏切りへの科罰」にコストがかかり、

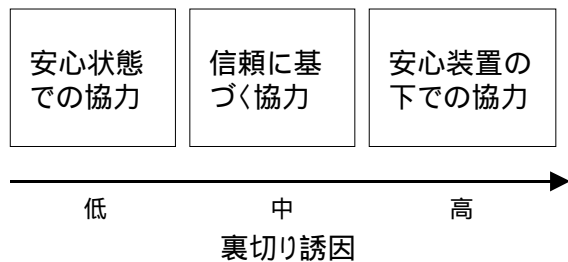


図3: 協力発生の概念図

しかも協力参入の利益を見送るという機会費用を負う。だが契約者が取引として請負人に報酬を与えれば、請負人と契約者との動機は両立する可能性がある。

こうした安心装置は、科罰の導入によって、より大きな報酬操作をエージェントに及ぼす。そこで、裏切り誘因が高い条件でも協力を確立できる可能性がある。

ここまでの筆者の推論をまとめたのが図3の概念図である。協力発生の様相は裏切り誘引の程度に応じて異なるだろう。まず、裏切り誘因が極めて低い場合は、人は特段の考慮もなく、安心して協力を結ぶことができる。裏切り誘因が高まるにしたがい、人は選択的に相手を選ぶ戦略に基づいて協力をするだろう。このとき重要になるのが信頼の高さである。しかしさらに裏切り誘因の強い状況では、権力のような安心装置を作ることによって人は協力を達成するだろう。

安心装置による協力の発生という以上のアイデアをテストするため次のシミュレーションを企画した。

### 3 方法

#### 3.1 基本設定

シミュレーションの基本設定は[3]と同じである。200のエージェントが社会を構成する。1世代は200ラウンドからなり、各ラウンドでランダムな順番でエージェントが協力呼びかけ人になる機会を得る。協力の呼びかけ/受諾はエージェントの戦略に基づく。協力関係の規模が大きいほど参加者の利得は大きい。しかし規模とともに呼びかけ人にはコストがかかるので、呼びかけ人にとり9名が「最適」規模である。裏切れば追加的利得が、他の参加者にはコストが生じる。戦略はビット列で表現する。信頼も戦略の一部である。世代内での利得合計の下位5%のエージェントの戦略は新たな戦略に入れ替わる。新たな戦略は上位5%のエージェントの戦略から利得に比例した確率で両親を

選び、交差をかけて定義する。世代の終了時に戦略の各次元には0.015の確率で突然変異が生じる。1試行で500世代まで計算した。

#### 3.2 安心装置の設定

出現した安心請負人に加入を申し出たエージェントはその請負人が主催する「安心クラブ」に属すると考える。請負人は所属料金の徴収と科罰に専念し、協力には参加しない。クラブ所属は各世代の開始直前に決まる。安心クラブは閉鎖的な集団と仮定する。つまり、成員が参加する協力は所属クラブに限定される。クラブ加入者は各ラウンドに、請負人が指定した額の利得を請負人に支払う。クラブ内で裏切りが生じたとき、請負人はその裏切りで得た利得の倍額の罰を裏切り者に科し、裏切り者をクラブから追放する。請負人に生じる罰のコストは、Axelrod[1]と同じく、罰の額の2/9である。

安心クラブ所属を決めるために戦略に次の次元を加えた：請負人となる申し出をするか否か、所属料金（8レベル）、安心クラブ所属を申し出るか否か、支払ってもよい所属料金上限（8レベル）。同料金で複数の請負人が申し出たとき、ランダムに1人に絞る。加入申込者は支払い上限内の料金を提示する請負人の中から1人をランダムに選んでそのクラブに加入する。加入者が2名より少ないクラブは消滅する。

#### 3.3 実験計画

シミュレーションを安心装置あり/なし×裏切り誘因要因（3水準）の6条件で実施した。安心装置なしの条件はすべて[3]と同じ設定である。裏切り誘因要因は低誘因条件（裏切り誘因係数  $w=0.001$ ）、中誘因条件 ( $w=0.5$ )、高誘因条件 ( $w=1.1$ )の3水準である。低誘因条件では裏切りはほとんど利益にならない。高誘因条件は社会的ディレンマの利得構造が保持される限界値に近く、[3]の最大誘因条件よりも誘因が高い。各条件で10試行を繰り返した。

### 4 シミュレーション結果

以下の分析はすべて、最後の100世代のデータに基づく。まず安心装置あり条件で安心装置は問題なく確立された。全体で85%のエージェントが安心クラブに属し、平均で6.3のクラブが供給されている。加

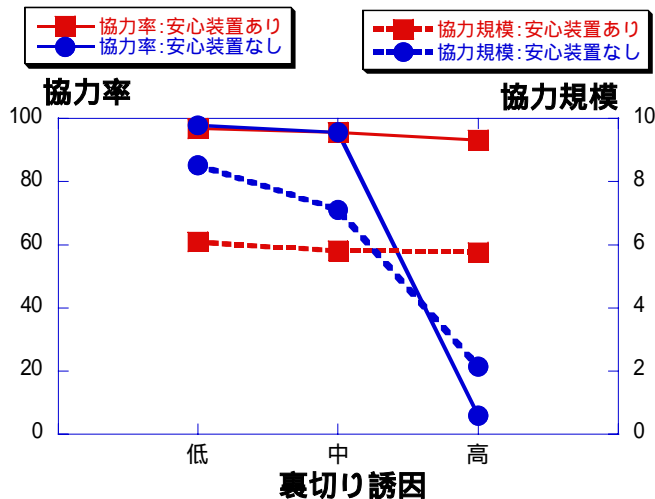


図4: シミュレーション結果  
(協力率と協力規模)

入料金は平均すると2番目に安い料金水準に相当する。

各条件での協力状況は図4に要約されている。まず安心装置なしの条件では、想定通り、低・中誘因条件で協力率と協力関係規模が高く、協力が確立されているのが分かる。同様に想定通り、高誘因条件では協力は生じていない。この結果は図2の結果の再現といえる。

他方、安心装置のある条件では、裏切り誘因の上昇とともに協力率と協力規模が若干下がるものの、ほぼ同水準で協力が維持されている。

安心装置なしの条件では信頼も図2の結果と符合している(図5)。低誘因条件、および協力が崩壊している高誘因条件で信頼は低く、中誘因条件で信頼が高まっている。安心装置ありの条件ではクラブ所属によって接触機会が限られるために安心装置なし条件と対応する信頼値を出すことはできない。しかし図5では、所属クラブ内での信頼、およびクラブ非所属者間での

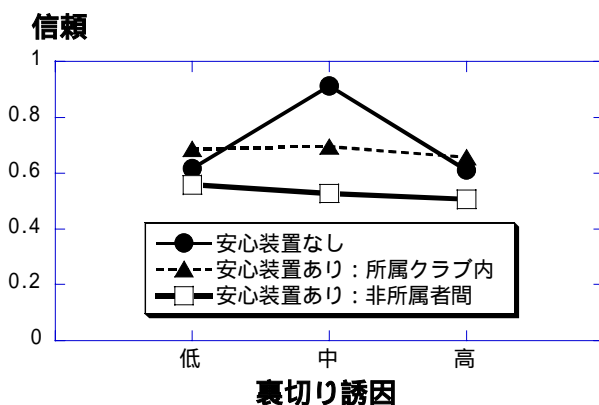


図5: シミュレーション結果(信頼)

信頼をプロットしてある。安心装置がある場合には、協力が生じているにもかかわらず、信頼は低水準であることが分かる。

## 5 考察

上記の結果は次のようにまとめることができる。

第1は、安心装置は、うまく確立されれば「水平的経路」では協力が達成できない(裏切り誘因が高い)状況でも協力を確立できることである。このことは裏切り誘因の高い状況、例えばヤクザの世界で、親分を戴き外部には閉鎖的・排他的な協力集団が形成されやすいことを意味するだろう。

第2は、安心装置のもとで協力が生じる場合には信頼は高まらないことである。図5は協力率が高い安心クラブの内部でも信頼は上がっていない。この結果は山岸[4]の信頼と安心の考察に別の解釈を与えるかも知れない。

このシミュレーション結果に筆者は限界と希望の両方を見出す。限界とはシミュレーションの設定に検討の余地が多いことである。結果は、安心クラブをどのように定義するか、安心装置の立ち上げ過程にどのような前提を置くかによって変わるだろう。このシミュレーションは図3に示したアイデアをデモンストレートするために最も単純な仮定で実施したものである。他方でこの結果の希望とは、例示した安心装置の出現が人間社会に普遍的な権力の創発経路を示唆することである。人間および人間社会の生存は協力を依存し、その協力を安定的に出現させるのが「料金を徴収する安心請負人」であるなら、この安心装置は容易に権力装置へと変貌することは意外ではない。

## 引用文献

- [1] Axelrod, R. (1986) An evolutionary approach to norms. *American Political Science Review*, 80, 1095-1111.
- [2] 高木英至 (2002) 協力呼びかけゲームの利得構造. 『埼玉大学紀要』, 38(2), 57-68.
- [3] Takagi, E. (2003) The evolution of inclusion mechanisms and trust in social dilemma situations. Paper presented at the 10<sup>th</sup> International Conference on Social Dilemmas, Marstrand, Sweden.
- [4] 山岸俊男 (1998) 『信頼の構造』, 東京大学出版会