

信頼がダメなら安心が協力を生み出す：

協力出現の計算機実験

高木英至（埼玉大学）

キーワード：安心、信頼、協力、central authority、進化シミュレーション

【要約】この報告は、central authority にあたる安心装置が進化的に出現可能であること、「信頼に基づく協力」が崩壊する裏切り誘因の強い状況でも安心装置は協力を出現させることができること、を計算機実験によって示す。

はじめに：協力はいかにして可能か？

協力は社会的ディレンマを内蔵するため、その出現は自明なことではない。筆者[3]は以前、n人協力関係の出現を検討するために「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た。このゲームでは、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。誰に呼びかけるか、呼びかけを受諾するか否かはエージェントの戦略による。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、呼びかけ人には「最適」規模が存在する。裏切れば追加的利得が、他の参加者にはコストが生じる([2])。

利得構造が社会的ディレンマとなる範囲で裏切り誘因を操作したときの主なシミュレーション結果は以下である。選択的受容戦略（相手の協力率を見て呼びかけ／受諾をする）が進化し、協力が確立される。

裏切り誘因が高い条件で信頼が高まる。ただし裏切り誘因が極端に高いと協力は崩壊し、信頼も低下する。

ここで生じた協力は「信頼に基づく協力」といえる。裏切り誘因がある程度高くなれば、エージェントは選択的受容戦略で裏切りから防御する以外にない。しかし信頼（未知の相手の協力確率の推定値）が低いと協力を形成できない。そこで選択的受容戦略と高信頼がペアになることで協力を生じさせる戦略が進化する。

しかし裏切り誘因がさらに高まるとき、信頼による協力は崩壊してしまう。この状況で信頼や specific trust を人為的に上げて効果はなかった。

解決策として考えられるのは、裏切りに対して制裁（罰）を科すような安心装置を設立して協力を確保す

ることである。安心装置は新たな利得要素を加えるため、信頼による協力が崩壊するほど高い裏切り誘因がある状況でも協力を確立する可能性がある。

しかし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、相互に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 のように、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。したがって誰かを権力者のような「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、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。

このアイデアを確認するために次のシミュレーションを実施した。

シミュレーションの設定

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、各試行で500世代を繰り返す。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。

安心装置を次のように設定した。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

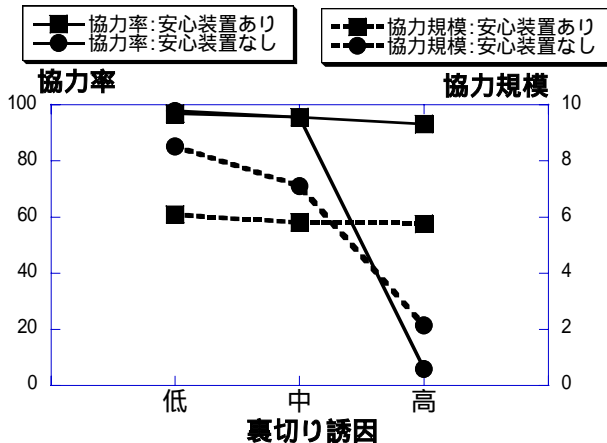


図1: シミュレーション結果(協力率と協力規模)

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。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で、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。

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 $\mu=0.001$ ）、中誘因条件（ $\mu=0.5$ ）、高誘因条件（ $\mu=1.1$ ）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。

結果と考察

最後の100世代のデータを分析する。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

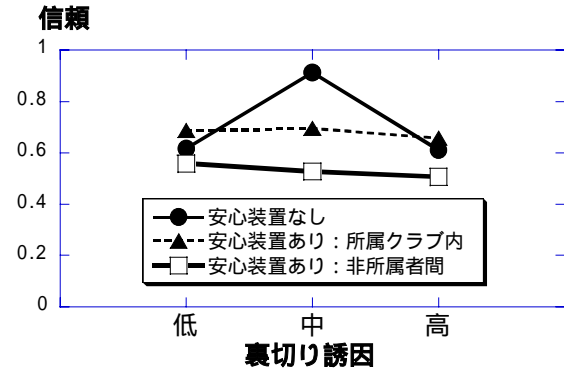


図2: シミュレーション結果(信頼)

のクラブが供給された。所属料金は平均すると2番目に安い料金水準に相当する。

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、協力が確立された（図1）。高誘因条件では協力は生じていない。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る。

安心装置なしの条件では、低誘因、および協力が崩壊している高誘因の場合で信頼は低く、中誘因条件で信頼が高まる（図2）。安心装置ありの条件については、所属クラブ内での信頼、およびクラブ非所属者間での信頼を図2で示す。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。

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、信頼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。この結果は山岸[4]の信頼と安心の考察に別の視点を与えるかも知れない。

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 10th 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