

強化学習主体からなる交換経済システム における貨幣の生成と崩壊

篠原 修二

概要

本研究では強化学習主体からなる交換経済システムのモデルを構築し、シミュレーションを行った。強化学習主体の学習能力は、割引率と呼ばれるパラメータによって特徴付けられる。小さい割引率は、主体が近視眼的であり即時的な報酬にのみ関心があることを意味する。一方割引率が大きいとき、主体は将来の報酬を現在の報酬と同等に評価する。シミュレーションの結果、ある特定のパラメータ領域においてのみ貨幣交換システムが出現することがわかった。割引率が小さいとき、主体は消費財にのみに価値を見出し、自分の生産財を直接消費財と交換しようとする物々交換戦略に固執する。割引率が増大するに伴って、貨幣が生成する。しかし、ある割引率がある値を越えると突然貨幣が崩壊し、互惠社会へと移行する。

Emergence and collapse of money in a trade system consisting of reinforcement learning agents

Shuji Shinohara

Abstract

In this study we explore whether money emerges in a multi-agent trade system in which agents evaluate goods based on their experiences. We have modeled artificially intelligent agents by applying reinforcement learning methods, and have made numerical experiments. Our simulation results show that the reinforcement learning agents develop a medium of exchange and come to acquire rewards more efficiently. Ability of the agents in our model is characterized by discount rate which determine how much they value future rewards at the present moment. Agent with small discount rate is concerned with only immediate reward, while one with large discount rate appreciates future and current rewards alike. The simulation results indicate that a barter economy emerges if the agents are myopic, a reciprocal community does if the agents are speculative. Monetary exchange system emerges between them.

1 はじめに

交換媒介機能は貨幣の重要な特性のひとつであり、これまで多くの研究者が交換媒体の生成過程を議論してきた。交換媒体としての貨幣の生成問題は、「それ自体としては何の有用性も持たないような財が、主体間の合意や法令などの強制力もなしに、なぜ／

どのようにして全ての主体に受け取られるようになるのか」という問題として定式化することができる。従来、経済学において貨幣の生成は、「欲望の二重の一致の困難」とその解消といった抽象で捉えられてきた。直接交換は、欲望の二重の一致の困難を伴う。つまり主体が直接交換を行うためには、自分の需要する財を所有しているだけでなく、同時に自分の

所有財も需要してくれるような取引相手を探し出さなければならない。主体の所有財や需要財が多様な財に分散している場合、主体がこのような取引相手を探し出すことは困難である。一方、交換経済に誰もが受け取る大域的財が存在するならば、その所有者は自分の欲する財の所有者を探し出すだけで交換を行うことができる。

メンガー [Men71] によれば、自分の利害のみを追求する利己的主体間の交換において、主体は自分の所有財をより販売力のある財と交換するようになる。このような行動は、自己触媒的に販売力のある財の販売力を更に高め、最終的に全ての主体がその財を受け取るようになる。メンガーに従った研究者達は、財の大域性を表す尺度として市場性という概念を導入した [KW89, 安富 00]。これらの研究によって、各主体が市場性を考慮して交換を行えば、経済内に貨幣が生成することが明らかにされた。しかしながら、貨幣の生成を扱う際に、市場性ような概念を前提とすることに対しては、論点先取あるいは遠近法的倒錯との批判がある [吉沢 93]。本研究では、市場性を前提とはせず、ただ学習能力のみをもつ主体を仮定する。ここでの学習とは、ある自発的な行動が起こり、それに引き続いて報酬が与えられるなら、その行動の強さは増大する、というものである。本稿では、このような学習能力をもつ主体からなる経済システムにおいて、どのように貨幣が生成するのかについて議論する。

2 モデル概要

本研究では、強化学習理論を用いてマルチエージェントシステムを構築しシミュレーションを行った。本節ではモデルの概要を述べる。

2.1 経済システム

まず、 $1, 2, \dots, N$ の番号を振られた N 人からなる主体集合 $A = \{1, 2, \dots, N\}$ と $1, 2, \dots, M$ という名前を付けられた M 種類の財集合 $G = \{1, 2, \dots, M\}$ からなる経済システムを考えよう。主体は、このシ

ステムにおいて財の生産、消費、交換を行う。各主体はある種の財を生産し、それとは別の種の財を消費することができる。本モデルでは簡単のため $N = M$ とし、主体 i は財 $i+1$ を生産し、財 i を消費することができるものとする。ただし主体 N は財 1 を生産し、財 N を消費するものとする。以上で定義したシステムでは、各主体が消費財のみに価値を見出し生産財と消費財を直接交換しようとしても、任意の主体間で交換は達成されない。

2.2 交換過程

交換の目的は、より多くの消費財を獲得することにある。主体は交換に先立って、経験に基づいて各財の価値を評価する。各主体は自らの価値評価に基づいて、所有財よりも価値の高い財を欲し、そのうちの一つを選び需要する。交換は、互いに相手の所有財を需要するならば、その時に限り成立する。

交換過程は、自分と交換可能な取引相手を探索することからはじまる。交換可能な取引相手とは、自分の欲する財を所有し、かつ自分の所有財を欲するような主体のことである。次に交換可能な取引相手の中から、最も価値の高い財を所有する主体を取引相手として選び出し、その主体と所有財を交換する。主体は、交換によって消費財を獲得するならば、それを消費し報酬 R を得ることができる。この時主体は、消費財を消費した後、直ちに生産財を生産し、それを所有財として次の交換に臨む。一方消費財以外の財を獲得した場合、報酬は得られない。その場合、主体は獲得した財をそのまま所有し、次の交換に持ち越す。主体は各時刻に一人ずつ順番に、このような交換過程を実行していく。ただし、その順番はランダムに決められる。 N 人全ての交換過程が一通り終了するまでの期間を 1 ターンと呼ぶ。すなわち、1 ターン = N ステップである。1 ターンが終了する度に、主体はランダムに並べ換えられ、またその順番に一人ずつ交換過程を実行する。これが繰り返されることで時間が進行する。なお、 TN ターンをもって 1 日とする。

3 強化学習と価値の形成

本モデルの主体は、経験に基づいて各財の価値を学習する。本モデルでは、その学習法として強化学習手法の一つである Temporal Difference 学習法を用いる。主体はある財を所有した際、その財が消費財と交換でき報酬を獲得できたならば、その財の価値を増加させる。主体 i が現時点での所有財 η_i に対して認める価値を $V_i(\eta_i)$ と表記するとすれば、交換過程実行後の新たな価値は以下の更新式で与えられる。

$$V_i(\eta_i) \leftarrow V_i(\eta_i) + \alpha[r_i + \gamma V_i(\eta'_i) - V_i(\eta_i)] \quad (1)$$

ただし η'_i は交換によって新たに獲得した財を表す。 α は、更新の度合を表すパラメータで学習率と呼ばれ、 $0.0 < \alpha \leq 1.0$ の値をとる。また γ は、割引率と呼ばれるパラメータで、 $0.0 \leq \gamma \leq 1.0$ である。割引率は、将来の報酬が現在においてどれだけの価値があるかを決定する。小さい割引率は、主体が近視眼的であり即時的な報酬のみに関心があることを意味する。一方割引率が大きいとき、主体は将来の報酬を現在の報酬と同等に評価する。ここで消費財については、主体が消費財を獲得したときには必ず消費するものと仮定し、その価値を R に固定する。また生産財についても、生産財を生産するためには一定の労働力 PC が必要であるものとし、その価値を PC に固定する。

4 シミュレーション結果

本節では、以上で定義したモデルのシミュレーション結果を示す。はじめにモデルの初期設定について述べる。各財の価値については、初期においては未知であるので、 $0 \sim 1.0$ の間の実数値をランダムに割り当てた。ただし生産財については、その財の生産に PC のコストがかかるものとし、その価値を PC とした。また消費財は必ず欲するものとし、その価値を R とした。生産財と消費財の価値は、時間に関して不変であると仮定する。またパラメータは、

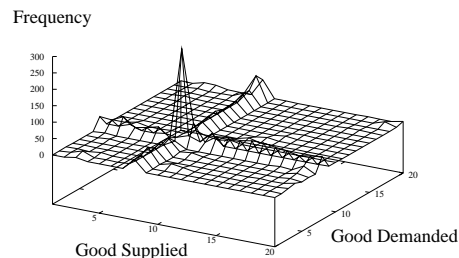


図 1: 需給状況 : $\gamma = 0.33$

$\alpha = 0.005$, $R = 1.0$, $N = M = 20$, $TN = 100$, $PC = 0.1$ に設定した。なお割引率 γ に関しては、その値を様々に変えて試行した。

図 1 は、割引率を $\gamma = 0.33$ に設定した場合のある日の経済の需要状況を示す。x 軸と y 軸は各々、主体が所有していた財、すなわち交換において供給しようとした財と需要した財を表す。また z 軸は、その日にそのような行動をとった主体総数を表す。図からわかるように、需給状況は財 8 を中心とした十字の構造を示す。これは、財 8 以外の様々な財を所有している主体は財 8 を需要し、財 8 の所有者は様々な財を需要していることを示している。なお、十字構造の中心にピークが見られるが、これは財 8 を所有し財 8 を需要する、すなわち財 8 を貯蓄する主体を表す。このように貨幣交換が生成する。

一方、図 2 は割引率を $\gamma = 0.2$ に設定した場合のある日の経済の需要状況を示す。図に示されたように、各主体は生産財を所有し消費財を需要するという物々交換戦略をとり続ける。しかしながら本モデルの設定では、欲望の二重の一致の困難のため交換は実現できない。このため獲得報酬は 0 となる。

一方、図 2 は割引率を $\gamma = 0.9$ に設定した場合のある日の経済の需要状況を示す。図に示されたように、各主体は様々な財を無秩序に需要する。

このように、割引率が小さく主体が即時的な報酬にのみ関心をもつとき、社会は物々交換経済となり、主体は効率よく消費財を獲得することができない。しかし割引率が大きくなり、主体が長期的観点から財を評価するようになるにつれて、貨幣交換システ

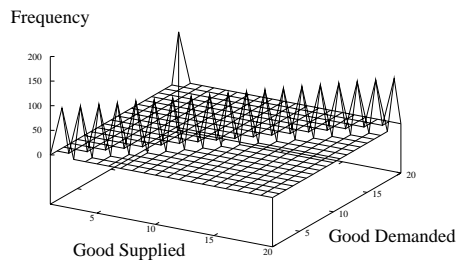


図 2: 需給状況 : $\gamma = 0.2$

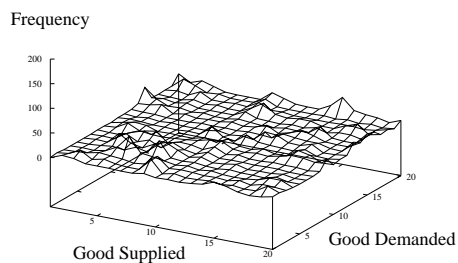


図 3: 需給状況 : $\gamma = 0.9$

ムが出現する。ただし貨幣交換が生成するパラメータ領域は、 $0.26 < \gamma < 0.4$ の領域に限定されており、 $0.4 \leq \gamma$ の領域においては貨幣が崩壊し、システムは互惠社会へ移行する。

参考文献

- [KW89] N. Kiyotaki and R. Wright. On money as a medium of exchange. *Journal of Political Economy*, Vol. 97, No. 4, pp. 927–954, 1989.
- [Men71] C. Menger. *Grundsätze der Volkswirtschaftslehre. Erster, allgemeiner Theil*. Wilhelm Braumüller, 1871.
- [安富 00] 安富歩. 貨幣の複雑性 生成と崩壊の論理. 創文社, 2000.
- [吉沢 93] 吉沢英成. 貨幣と象徴. ちくま学芸文庫, 1993.