

# 物々交換経済と互惠社会の間に出現する貨幣交換システム

篠原 修 二<sup>†,††</sup>

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、数値実験を行った。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、主体が即時的な報酬のみに関心を持つことを意味する。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。数値実験の結果、ある特定のパラメータ領域においてのみ貨幣が生成することが分かった。割引率が小さいとき、主体は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従う。一方割引率が大きいとき、他主体に生産財を与える代わりに、他主体から消費財を受けられるという互惠社会が成立する。貨幣交換システムは両者の中間領域において出現する。

## A Monetary Exchange System Emerging between a Barter Economy and a Reciprocal Community

SHUJI SHINOHARA<sup>†,††</sup>

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 a reinforcement learning method, 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, whereas a reciprocal community does if the agents are speculative. Monetary exchange system emerges between them.

### 1. はじめに

従来、交換媒体としての貨幣の生成は、欲望の二重の一致の困難とその解消といった抽象でとらえられてきた。たとえば、権力者が一元的に財を管理し分配するような共同体社会や、各成員が他の成員に自分の所有財を無償で提供するような互惠社会においては、人々は必要に応じて効率良く財を入手できただろう。しかしながら、社会が大規模になり、自律分散した個人が各自の利害に基づいて交換を行うようになるにつれて、欲望の二重の一致の困難が発生する。つまり、このような主体どうし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はなく、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。各主体の所有財や需要財が多様な財に分散している場合、このような取引相手を探し出すことは困難である。一方、経済内に誰もが受け取る貨幣のような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。このような経済では、直接交換を行うよりもいったん自分の所有財を貨幣と交換しておき、その後貨幣と需要財を交換する、つまり貨幣を媒介とした間接交換を行う方が、効率良く需要財を獲得できる。このように貨幣の生成は、財の分配を特定の個人や組織の権力に委ねることなく、また各主体の利他的な意識に頼ることもなく、社会全体の効率を上げることを可能とする。

各主体の欲望の多様性にもかかわらず、貨幣はなぜ/どのようにしてすべての主体に受け取られるようになるのだろうか。それは、貨幣法制説が主張するように、国家権力などによってそうすることを義務付けられたためかもしれないし、貨幣商品説が主張するように、

<sup>†</sup> 京都産業大学

Kyoto Sangyo University

<sup>††</sup> 日本学術振興会特別研究員

JSPS Research Fellow

現在、株式会社デジタルドリーム

Presently with Digitaldream Inc.

数ある財の中から自然発生的に貨幣のような財が生成したのかもしれない。ただしここで注意したいのは、はじめにどのような理由によって貨幣が生成したにせよ、貨幣が貨幣としての機能を維持していくためには、経済内のすべての主体が一致して、貨幣を使用し続ける必要があるということである。自分1人がある財を貨幣として受け取るとしても、望ましい結果は得られない。この意味で、貨幣は1つの慣習といえるだろう。

Mengerによれば、自分の利害のみを追求する利己の主体間の交換において、各主体は自分の所有財をより販売力のある財、つまり交換しやすい財と交換するようになる<sup>10)</sup>。このような行動は、最も販売力のある財の販売力をさらに高め、最終的にすべての主体がその財を受け取るようになる。Mengerに従った研究者たちは、財の販売力の度合いを表す尺度として市場性という概念を導入した。たとえばKiyotakiとWrightは、財の市場性を、ある財を所有しある財を消費対象とする主体が経済内にどのような割合で存在するのかわる確率分布の形で導入した<sup>8)</sup>。この研究によって、各主体が市場性を考慮して交換を行うような経済においては、ある条件のもとで貨幣が生成しうることが明らかになった。ただしこの研究は静的な均衡分析が中心であり、貨幣の生成条件を明らかにしたものの、均衡が達成される過程、すなわち貨幣の生成過程それ自体を扱っているとはいいい難い。このため最近では、Kiyotaki-Wrightモデルの環境設定のもとで、実際の人間を使った経済実験<sup>3),4)</sup>やマルチエージェントシステムを用いた数値実験<sup>9)</sup>などが行われている。

安富は、主体に「他人の受け取るものを受け取る」という戦略を実装することによって、より直接的に市場性概念を導入した<sup>20)</sup>。Kiyotaki-Wrightモデルにおいて各主体に与えられる市場性の情報は、すべての主体において共通であると仮定されている。これに対し安富は、他主体との知識交換を通して市場性に関する情報を蓄積していく主体をモデル化した。彼は、これに加えてノイズや淘汰も実装した独自の進化的モデルを構築し、貨幣の生成過程だけでなく崩壊過程についても論じている。

しかしながら、貨幣の生成過程を扱う際に、市場性の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<sup>13),21)</sup>。たとえば吉沢は以下のように述べている。

じつは「販売可能度」といった概念は一般的な交換手段のもとで存在しうるので、その概念自身、貨幣を含意しているのである。商品それぞれに「市場性」の度を前提する、すなわち「販売可能度」という概念を前提することは、形成さるべき貨幣を前提することにも等しい(吉沢(1993) p.117)。

本稿では市場性を明示的に考慮することはせず、単純な学習能力のみを有する主体を仮定する。ここでの学習とは、「成功や報酬は報酬を受ける行動の学習を進め、罰や失敗はその行動を反復する傾向を減少させる」というものである。これは、心理学の分野では効果の法則と呼ばれ、様々な動物を用いて古くから研究されてきた<sup>2),15)</sup>。また工学の分野においても、このような学習は、強化学習過程としてモデル化されてきた<sup>14),19)</sup>。本モデルでは、経験に基づいてより多くの報酬、すなわち消費財を獲得することを学習する主体を、強化学習手法を用いてモデル化する。ここで主体の学習能力は、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。本稿の目的は、このような学習能力を持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わるを明らかにすることである。

## 2. 経済システム

本章では、モデルで扱う経済システムの枠組みを定義する。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。簡単のために、主体は各時点において1種類1単位の財を所有し、交換も1度に1種類1単位の財のみが許されるものとしよう。また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。すなわち、自分の生産財を、自分で消費することはできない。主体は、交換によって消費財を獲得したときには、それを消費し報酬  $R$  を得ることが

「一方の貨幣商品説とは貨幣とはそれ自体が価値をもつ商品その起源とし、ひとつのあいだの交換活動のなかから自然発生的に一般的な等価物あるいは一般的な交換手段へと転化したという主張である。他方の貨幣法制説とは、貨幣とはそれ自体が商品としての価値をもつ必要はなく、共同体の申し合わせや皇帝や君主の勅令や市民の社会契約や国家の立法にその起源をもとめることができるという主張である」(岩井(1993) p.81)。

人間以外で市場性のような概念を持つ生物が存在するかどうかは定かではない。しかしこのように単純な学習能力ならば、ラットやハトをはじめ様々な動物において認められている。

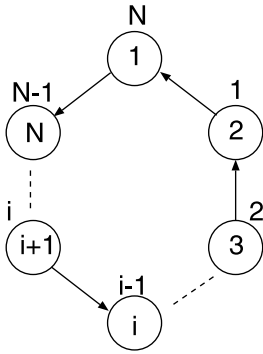


図1 経済システム．丸上の数字と丸で囲まれた数字は、各々主体番号とその主体の生産財を表す．また矢印は、矢印の始点で表される主体が矢印の終点で表される主体の生産財を消費財とすることを表す．この状況ですべての主体が消費財のみを欲するならば、任意の主体間で交換は実現されない．いわゆる欲望の二重の一致の困難である．

Fig. 1 Trade system.

できる．消費財を消費した主体は、ただちに生産財を生産し、それを次の所有財とする．ここで、ある時点における主体  $i$  の所有財を、 $\eta_i \in \mathbf{G}$  と表記する．また、主体  $i$  の消費財と生産財を各々、 $\theta_i \in \mathbf{G}$ 、 $\pi_i \in \mathbf{G}$  と表記する．本モデルでは簡単のため  $N = M$  とし、また主体  $i$  は財  $i$  を消費し、財  $i+1$  を生産することができるものとする．すなわち、 $\theta_i = i$ 、 $\pi_i = i+1$  である（図1）．ただし主体  $N$  は財  $N$  を消費し、財  $1$  を生産するものとする（ $\theta_N = N$ 、 $\pi_N = 1$ ）．なお、各主体の生産財と消費財は、時間に関して不変であると仮定する．

### 3. 交換過程

本章では、交換過程について述べる．交換の目的は、より多くの消費財を獲得することである．交換に先立って、主体は自らの経験に基づいて各財の価値を評価する．ここで、主体が各財の価値を評価するために用いる価値関数を導入し、ある時点において主体  $i$  がもつ価値関数を  $V_i: \mathbf{G} \rightarrow \mathbb{R}$  と表記する．ただし、 $\mathbb{R}$  は実数集合であり、価値関数とは各財に対して実数で表される価値を割り当てる関数である．主体は、他主体との相互作用を通して、各財の価値を学習していく．その学習法については、5章で述べる．

主体は自らの価値関数に従って、なるべく価値の高い財を入手しようとする．はじめに主体（ $i$  と表記する）は、自分が最も高い価値を認める財の所有者のもとへ赴き交換を提案する．このとき相手が交換の要求に応じてくれるならば、取引が成立し主体  $i$  の交換過程は終了する．ただし交換を提案された主体は、主体

$i$  の所有財に自分の所有財よりも高い価値を認めるならば取引に応じるものとする．一方交換を拒否された場合、 $i$  は2番目に高い価値を持つ財の所有者に交換を提案する．ここでも交換を拒否された場合、 $i$  はさらに3番目に高い価値を持つ財の所有者に交換を提案する．以下同様に、主体  $i$  は交換が実現できるか、あるいは自分の所有財よりも高い価値を持つ財の所有者が存在しなくなるまで、交換の提案を繰り返す．

このような交換過程をモデル化するために、まずある時点において主体  $i$  と交換可能な相手からなる集合  $O_i$  を考える．主体  $i$  と交換可能な取引相手とは、 $i$  が欲する財を所有し、かつ  $i$  の所有財を欲してくれるような主体のことである．つまり、互いが相手の所有財に自分の所有財よりも高い価値を認めるならば、彼らは交換可能である．以下に  $O_i$  の定義を示す．

$$O_i = \{j | j \in \mathbf{A}, V_i(\eta_j) < V_i(\eta_i), V_j(\eta_i) > V_j(\eta_j)\} \quad (1)$$

次に主体  $i$  は、自分と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選び出す．つまり、ある時点における主体  $i$  の取引相手を、 $o_i$  と表記することにすれば、

$$o_i = \begin{cases} j \in O_i, s.t. V_i(\eta_j) = \max_{k \in O_i} V_i(\eta_k) & O_i \neq \emptyset \text{ のとき} \\ 0 & \text{それ以外} \end{cases} \quad (2)$$

である．ただし、ここで0は経済内に取引相手が存在しないことを表す．なお、最も価値の高い財を所有する主体が複数存在する場合は、それらの主体のうち、主体  $i$  の所有財に最も高い価値を見出す主体を取引相手とする．

取引相手が決定したならば、 $i$  はその相手の所有財を需要する．もし交換可能な取引相手が存在しないならば、そもそも交換は不可能である．このとき  $i$  は、何を需要しても、あるいは何も需要しなくても結果は同じであるが、便宜上その時点で経済内に存在する財のうち、最も価値の高い財を需要することにする．取引相手が存在しない状況として、以下の2つの場合が考えられる．第1は、自分の所有財より価値の高い財が経済内に存在するが、その所有者が交換に応じてくれないような場合である．このとき  $i$  は交換しなくても交換できない．第2はそもそも欲する財が経済内に存在しない、すなわち自分の所有財が経済内で最も価値の高い財である場合である．このような場合、 $i$  は交換できないのではなく、交換しようとしていないのである．本稿では、このような場合、主体  $i$  は自分の所有

財を貯蓄していると見なすことにする．整理すると，ある時点における主体  $i$  の需要財  $\xi_i$  は，以下のように記述できる．

$$\xi_i = \begin{cases} \eta_{o_i} & o_i \neq 0 \text{ のとき} \\ \eta_j, s.t. V_i(\eta_j) = \max_{k \in A} V_i(\eta_k) & \text{それ以外} \end{cases} \quad (3)$$

取引相手と需要財が決まったならば，主体  $i$  は取引相手  $o_i$  と互いの所有財を交換する．もし獲得した財が消費財ならば，それを消費し報酬  $R$  を得ることができる．そうでないなら，獲得報酬は 0 である．消費財を消費した主体はただ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主体は，獲得した財をそのまま所有財とし次の交換に臨む．交換可能な相手が存在しない場合，この主体は交換を行うことができない．その場合は，現在の所有財をそのまま次の交換に持ち越す．つまり次時点における主体  $i$  の所有財を  $\eta'_i$  と表記することにすれば，

$$\eta'_i = \begin{cases} \pi_i & o_i \neq 0 \text{ かつ } \xi_i = \theta_i \text{ のとき} \\ \xi_i & o_i \neq 0 \text{ かつ } \xi_i \neq \theta_i \text{ のとき} \\ \eta_i & o_i = 0 \text{ のとき} \end{cases} \quad (4)$$

である．また，状態が  $\eta_i$  から  $\eta'_i$  に移行する際に，主体  $i$  が獲得する報酬  $r_i$  は，以下のように計算される．

$$r_i = \begin{cases} R & o_i \neq 0 \text{ かつ } \xi_i = \theta_i \text{ のとき} \\ 0 & \text{それ以外} \end{cases} \quad (5)$$

なお，主体  $i$  に取引相手が存在するとき，その相手  $o_i$  の所有財と報酬も同様に更新される．

以上をもって，主体  $i$  の交換過程は終了する．主体は各時刻に 1 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． $N$  人すべての交換過程がひととお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 =  $N$  時間ステップである．1 ターンが終了するたびに，主体はランダムに並べ換えられ，またその順番に 1 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．すなわち，1 日 =  $TN \times N$  時間ステップである．

#### 4. 欲望の二重の一致の困難とその解消

図 1 に示したように，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．ここで注意すべきは，経済全体で見れば，

すべての財に関してその需要量と供給量が一致しているにもかかわらず，交換は実現されないということである．この点を明確にしよう．

まず，ある時点において財  $\eta$  を所有し財  $\xi$  を需要している主体数  $\Phi(\eta | \xi)$  を，以下のように表す．

$$\Phi(\eta | \xi) = |S_\eta \cap D_\xi| \quad (6)$$

ただし， $S_\eta, D_\xi$  は，各々

$$S_\eta = \{i | i \in A, \eta = \eta_i\} \subseteq A \quad (7)$$

$$D_\xi = \{i | i \in A, \xi = \xi_i\} \subseteq A \quad (8)$$

である．すなわち  $S_\eta(D_\xi)$  とは，その時点において財  $\eta$  を所有する（財  $\xi$  を需要する）主体からなる集合である．なお， $X$  が集合であるとき， $|X|$  は集合  $X$  の要素数を表す．

次に，経済全体の需給量バランスを表す尺度として以下のような変数  $Consv$  を導入する．

$$\begin{aligned} Consv &= \sum_{\eta=1}^N \left| \sum_{\xi=1}^N \Phi(\eta | \xi) - \sum_{\xi=1}^N \Phi(\xi | \eta) \right| \\ &= \sum_{\eta=1}^N \left| \sum_{\xi=1}^N (\Phi(\eta | \xi) - \Phi(\xi | \eta)) \right| \end{aligned} \quad (9)$$

条件  $Consv = 0$  は， $\forall \eta \in G$  に対して，経済全体で財  $\eta$  を所有する主体数と財  $\eta$  を需要する主体数が等しいことを表す．さらに，経済の対称性を表す尺度として次のような変数  $Sym$  を導入する．

$$Sym = \sum_{\eta=1}^N \left( \sum_{\xi=1}^N |\Phi(\eta | \xi) - \Phi(\xi | \eta)| \right) \quad (10)$$

条件  $Sym = 0$  は， $\forall \eta, \xi \in G$  に対して，財  $\eta$  を所有し財  $\xi$  を需要する主体数と，財  $\xi$  を所有し財  $\eta$  を需要する主体数が等しいことを表す．式 (9) と (10) を比較すれば，必ず  $Consv \leq Sym$  が成立する，つまり  $Sym = 0$  の方がより厳しい条件であることが分かる．たとえば 図 1 の状況で，各主体が生産財を所有し消費財を需要するとすれば， $Consv = 0$  であるが， $Sym = N$  となる．

ここで，2 つの条件  $Consv = 0$  と  $Sym = 0$  の意味について考えよう．まず  $Consv = 0$  は上述のように，経済全体で所有されている財量と需要されている財量が，財ごとにつりあっていることを示している．これは，いったん各主体の所有財を 1 か所に集め，それから各主体の需要に応じて財を分配するならば，ちょうど過不足なくすべての主体の需要を満足させることができることを意味する．特定の個人や組織などの権力が不在の状況で，このような財の分配を実現しうる社会としては，各主体が他主体に自分の生産財を無償で提供するような互惠社会があげられよう．一方

$Sym = 0$  が成立するときには、自律分散した主体間の交換のみによって、すべての主体の需要を満足させることができる。

次に貨幣交換が成立するための条件を考えよう。ここで、貨幣を  $M$  と表記する。また  $M$  が大域的であるとは、以下のような条件が満たされていることとする。

- 交換を行おうとするすべての主体は、少なくとも  $M$  を所有しているか、 $M$  を必要しなければならない。

この条件は、貨幣交換経済においては任意の財は貨幣との間でのみ交換されうること、よって貨幣以外の財を所有する主体はいったん貨幣と所有財を交換して（売る）、その後貨幣を何らかの財と交換する（買う）ことを考えるならば妥当な条件であろう。またこの条件は、Jones<sup>6)</sup> による交換媒体の定義にも合致する。この条件は、次のように記述することができる。

$$|S_M \cup D_M| = \left| \bigcup_{\xi \in G} D_\xi \right| \quad (11)$$

左辺は、 $M$  を所有するかあるいは需要する主体数を表す。一方右辺は、交換を行おうとする主体、すなわち何らかの財を需要する主体の数を表す。本モデルでは、主体は 1 種類 1 単位の財のみの所有と需要を許されているので、条件 (11) の両辺は、 $\Phi$  を用いて以下のように書き換えられる。

$$\begin{aligned} \sum_{\xi=1}^N \Phi(M | \xi) + \sum_{\eta=1}^N \Phi(\eta | M) - \Phi(M | M) \\ = \sum_{\xi=1}^N \sum_{\eta=1}^N \Phi(\eta | \xi) \end{aligned} \quad (12)$$

ここで条件 (12) の制約のもとで、2 つの条件  $Consv = 0$  と  $Sym = 0$  が同等であることを示す。つねに  $0 \leq Consv \leq Sym$  が成立するので、 $Sym = 0$  ならば  $Consv = 0$  は明らかである。よってここでは、条件 (12) を仮定したときに、 $Consv = 0$  ならば  $Sym = 0$  であることを証明する。式 (9), (10) より、 $Consv = 0$  と  $Sym = 0$  は各々次のように書き換えられる。

$$\forall \eta, \sum_{\xi=1}^N \Phi(\eta | \xi) = \sum_{\xi=1}^N \Phi(\xi | \eta) \quad (13)$$

$$\forall \eta, \xi, \Phi(\eta | \xi) = \Phi(\xi | \eta) \quad (14)$$

また、条件 (12) より、次のことが導かれる。

$$\forall \eta, \xi(\eta, \xi \neq M), \Phi(\eta | \xi) = 0 \quad (15)$$

したがって、式 (13), (15) を仮定したとき、式 (14) がいえればよい。以下にその証明を示す。

- $\eta, \xi \neq M$  のとき

式 (15) より、 $\Phi(\eta | \xi) = 0, \Phi(\xi | \eta) = 0$  である。したがって  $\Phi(\eta | \xi) = \Phi(\xi | \eta)$ 。

- $\eta, \xi = M$  のとき

$\Phi(M | M) = \Phi(M | M)$  より明らかである。

- $\eta \neq M, \xi = M$  のとき

式 (15) より、 $\forall \zeta(\zeta \neq M)$  に対して  $\Phi(\eta | \zeta) = 0$  であるので、 $\Phi(\eta | M) = \sum_{\zeta=1}^N \Phi(\eta | \zeta)$ 。同様に、 $\forall \zeta(\zeta \neq M)$  に対して  $\Phi(\zeta | \eta) = 0$  であるので、 $\Phi(M | \eta) = \sum_{\zeta=1}^N \Phi(\zeta | \eta)$ 。また、式 (13) より、 $\forall \eta$  に対して  $\sum_{\zeta=1}^N \Phi(\eta | \zeta) = \sum_{\zeta=1}^N \Phi(\zeta | \eta)$ 。したがって、 $\Phi(\eta | M) = \Phi(M | \eta)$ 。

- $\eta = M, \xi \neq M$  のとき

$\eta \neq M, \xi = M$  のときと同様。

よって、 $\forall \eta, \xi$  に対して  $\Phi(\eta | \xi) = \Phi(\xi | \eta)$  である。

以上に示されたように、条件 (12) が成立するような経済では、条件  $Consv = 0$  が成立するならば、必ず条件  $Sym = 0$  が成立する。つまり、貨幣のような大域的財が存在する状況のもとでは、財を 1 つに集め、各主体の需要に応じて分配することと、各主体が欲望の二重の一致に基づいて交換を行うことの区別が無効にされる。このことは、 $Sym = 0$  と  $Consv = 0$  との間のギャップ、すなわち欲望の二重の一致の困難は、貨幣の導入によって解消されるということを示している。

本モデルの枠組みで条件 (12) が満たされる可能性があるのは、すべての主体が各自の消費財に最高の価値を、また貨幣  $M$  に 2 番目に高い価値を見出し、それ以外の財を無価値と見なす場合である。消費財を獲得した主体は、ただちにそれを消費し生産財と置き換えるので、消費財を所有している主体は存在しない。すなわち、すべての主体は、貨幣があるいは自分にとって無価値な財を所有する。主体は、自分の生産財を無価値な財と交換することはないので、結局すべての主体は、貨幣が生産財を所有する。貨幣を所有する主体は、貨幣より価値の高い唯一の財である消費財を欲する。もし彼の消費財を所有する主体が存在するならば、その相手は必ず貨幣を欲するのでこの貨幣の所有者は消費財を獲得することができる。その後彼は、消費財を消費し生産財を生産するので生産財の所有者となる。一方生産財を所有する主体は、それより価値の高い財である貨幣と消費財を欲する。この主体は、彼の消費財を所有し、かつ彼の生産財を消費財とするような主体、すなわち物々交換可能な取引相手が存在すれば、消費財を需要する。しかし本モデルの枠組みでは、そのような取引相手はそもそも存在しない。これに対し、彼の生産財を消費財とするような貨幣の所

有者ならば存在する可能性がある。その場合、彼は貨幣を需要する。

たとえば図 1 の状況で、財 1 を貨幣としよう。すなわち、すべての主体が財 1 に高い価値を認めているとする。また簡単のために主体は、主体 N から主体番号の逆の順番に交換過程を実行していくとしよう。まず財 1 の所有者である主体 N は自分の消費財である財 N を欲する。財 N の所有者である主体 N-1 にとって財 1 は消費財ではないけれども、貨幣であるためにそれを需要する。このとき主体 N と N-1 の間で交換が成立し、主体 N は消費財 N を入手できる。次に財 1 の所有者となった主体 N-1 は、自分の消費財である財 N-1 を欲する。財 N-1 の所有者である主体 N-2 も同様に貨幣である財 1 を需要するので、主体 N-1 と N-2 の間で交換が成立する。以下同様に主体  $i$  と主体  $i-1$  の間で交換が成立し、財 1 は反時計回りに主体間を還流する。この場合、経済内で行われるすべての交換において財 1 が使用されることになり、条件 (12) が満たされる。

## 5. 強化学習と価値の形成

4 章で示したように、本モデルの枠組みでは、各主体が消費財のみに価値を見出しそのみを需要し続けるならば、欲望の二重の一致の困難のために交換は実現されない。一方、経済内にすべての交換において使用される大域的財、すなわち貨幣が存在するならば、その困難は解消される。もし貨幣の価値が所与であり、あらかじめ各主体がその価値を知っているならば、その価値を根拠として貨幣を欲すればよい。しかしながら、財の価値は外的に与えられるものではなく、主体間の相互作用の結果として内生的に形成されるものである。つまり、初期状態では、各財の価値は未知であると同時に未定である。主体は他主体との交換を通して、各財の価値を学習（形成）していかなければならない。本稿では、このような過程をモデル化するために、強化学習と呼ばれる手法を用いる。

強化学習では、主体の状態と各状態に対して価値を割り当てる価値関数を定義する。主体は、価値関数に基づいてある行動を選択する。この選択方法のことを

方策と呼ぶ。つまり主体は、ある状態のもとである方策に従いある行動をとる。その際、環境からある報酬を受け取り次の状態に移行する。ここである状態の価値とは、ある状態においてその後ある方策をとっていったときに最終的に得られる総報酬の期待値のことである。強化学習主体の唯一の目的は、最終的に得られる総報酬を最大化することである。

さて強化学習は、状態の価値の学習と、価値に従った行動の選択という 2 つの部分から構成される。たとえば強化学習の中にはモンテカルロ法という手法があるが、これはまず試行錯誤によって各状態の価値を学習し、次にそれらの価値に基づいて何らかの行動をとるというものである。しかし、この手法では学習が完了するまで、何の行動をとることもできない。これに対し TD (Temporal Difference) 学習と呼ばれる手法では、主体が各時刻において価値の推定値を利用することで、学習しつつ何らかの行動をとることを可能にする。

ここで、TD 学習の中でも最も単純な TD(0) と呼ばれる学習法について説明する。詳しくは文献 19) を参照されたい。ある時刻  $t$  におけるある状態  $s_t$  の価値とは、主体がある方策に従って行動したとき、その時点から将来にわたって獲得できる期待収益のことである。時刻  $t$  の後に得られた報酬の系列を、 $r_{t+1}, r_{t+2}, r_{t+3}, \dots$  と表すなら、それらの報酬の合計である収益  $R_t$  は次式で与えられる。

3 章で示したように、本稿のモデルでは、行動選択の方法として最も価値の高い財を需要するという方策を用いている。これはグリーディ方策と呼ばれる<sup>19)</sup>。

ここで、価値と報酬の違いについて述べたい。たとえば食べ物のように、主体に直接的な効用を与える対象、すなわち報酬は、その主体にとって価値が高いといえる。しかし、価値の高い対象が主体に直接的な効用をもたらすとは限らない。たとえば、以下のような学習実験を考える。まず被験体である主体 A に対して、X と Y という 2 つの対象を選択肢として与える。ただし、X と Y は A に対して何の効用ももたらさないとする。つまりどちらを選択しても、主体がそのことによって得られる効用は 0 である。このような場合、A がどちらの対象を選択するかはランダムであり、X と Y を選択する頻度に有意な差は生じないと考えられる。次に、A が X を選択したときには直後に報酬 (たとえば食べ物) を与え、一方 Y を選択したときには何も与えないような状況を考える。この実験においては、試行回数の増加とともに、X を選択する頻度が増大すると予想される。このとき、X の方が Y よりも価値が高いと見なすことができる。つまり学習の結果、A にとって X の価値は増大する。ただし、学習の事後においても、X が食べ物に変化するわけではなく、X それ自身が A に与える効用は 0 である。このように本稿では、価値と報酬を区別する。ここで重要な点は、対象の価値は状況に依存するという点である。たとえば学習実験を、Y を選択したときのみ報酬が与えられるという状況に変更したとすれば、主体 A にとって価値が高いのは X ではなく Y の方である。

ここで強化学習手法を用いる理由としては、第 1 に実装が簡単であること、第 2 に強化学習手法では価値関数が明示的に導入されること、第 3 に実際の人間と強化学習主体の行動を比較した研究があるということなどがあげられる。最後の点に関して、Iwasaki ら<sup>5)</sup> は公共財ゲームにおいて、実際の人を使った実験結果と強化学習主体によるシミュレーション結果を比較し、強化学習手法によって、人の協調行動パターンが再現できることを示している。

$$R_t = r_{t+1} + r_{t+2} + r_{t+3} + \dots + r_T \quad (16)$$

ここで  $T$  は最終時間ステップである．なお，初期状態から  $T$  までをエピソードと呼ぶ．式 (16) では，収益が将来に獲得しうる報酬の単純和で表されているが， $T = \infty$  で  $R_t$  が発散するような場合には，以下のよう割引率を導入した収益を考えることもできる．

$$\begin{aligned} R_t &= r_{t+1} + \gamma r_{t+2} + \gamma^2 r_{t+3} + \dots + \gamma^{T-(t+1)} r_T \\ &= \sum_{k=1}^{T-t} \gamma^{k-1} r_{t+k} \end{aligned} \quad (17)$$

ここで  $\gamma$  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． $\gamma = 1.0$  のときは，式 (17) は式 (16) と等価であり，将来の報酬を現在の報酬と同等に評価する．つまり式 (16) は式 (17) の特殊な場合と考えられる．一方， $\gamma = 0.0$  のときは， $R_t = r_{t+1}$  であり主体は即時的な報酬のみに関心があることを意味する．

ここで状態  $s_t$  の価値の推定値を  $V(s_t)$  と表記しよう．経験に基づいて価値を推定する最も単純な方法は，過去のエピソードで状態  $s_t$  となったときに，その後実際に得られた収益を平均することである．そのような場合が過去に  $k$  回あったとし，その平均が  $V(s_t)$  であったとしよう．今回のエピソードで再び状態  $s_t$  が生起し，その後得られた収益が  $R_t$  だったとすれば，新たな推定値  $V'(s_t)$  は以下のように与えられる．

$$\begin{aligned} V'(s_t) &= \frac{1}{k+1} (R_t + kV(s_t)) \\ &= \frac{1}{k+1} (R_t + kV(s_t) + V(s_t) - V(s_t)) \\ &= \frac{1}{k+1} (R_t + (k+1)V(s_t) - V(s_t)) \\ &= V(s_t) + \frac{1}{k+1} (R_t - V(s_t)) \end{aligned} \quad (18)$$

ここで  $1/(k+1) = \alpha$  とおくと， $V(s_t)$  の更新式は以下のように与えられる．

$$V(s_t) \leftarrow V(s_t) + \alpha [R_t - V(s_t)] \quad (19)$$

$\alpha$  は，価値関数の更新の割合を表すパラメータで学習率と呼ばれる． $\alpha$  は  $k$  の増大ともなまって 0 に近づくが，定数として扱われることも多い．これは過去の情報よりも最近の情報に，より大きい重みを与えることを意味する．本稿のモデルでも  $\alpha$  は定数値とする．ただし， $0.0 < \alpha \leq 1.0$  である．

さて，式 (19) はモンテカルロ法と呼ばれるものであるが， $R_t$  はエピソードが終了するまで確定しない．ここで，価値を学習しつつ同時に何らかの行動をとることができるように，式 (19) を以下のように変更する．

式 (17) は，

$$\begin{aligned} R_t &= \sum_{k=1}^{T-t} \gamma^{k-1} r_{t+k} \\ &= r_{t+1} + \gamma \sum_{k=1}^{T-(t+1)} \gamma^{k-1} r_{(t+1)+k} \end{aligned} \quad (20)$$

と書き換えられるが，ここで  $\sum_{k=1}^{T-(t+1)} \gamma^{k-1} r_{(t+1)+k}$  の推定値として次の状態の推定価値  $V(s_{t+1})$  を利用する．このとき，式 (19) は次のように変更される． $V(s_t) \leftarrow V(s_t) + \alpha [r_{t+1} + \gamma V(s_{t+1}) - V(s_t)]$  (21) これが，TD(0) と呼ばれるものである．本モデルでは，価値関数の学習法として以上で紹介した TD(0) を用いる．ただし主体の状態は，その主体の所有財をもって表現する．ここで，ある時点における主体  $i$  の所有財を  $\eta_i$ ，次状態を  $\eta'_i$ ，また状態が  $\eta_i$  から  $\eta'_i$  に移行する際に主体  $i$  が獲得する報酬を  $r_i$  と表記するとすれば，主体  $i$  の持つ価値関数  $V_i$  の更新式は以下のように与えられる．

$$V_i(\eta_i) \leftarrow V_i(\eta_i) + \alpha [r_i + \gamma V_i(\eta'_i) - V_i(\eta_i)] \quad (22)$$

ただし消費財については，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．したがって式 (22) を用いるのは，所有財が消費財と生産財以外の場合に限る．このため， $r_i$  は 0.0 の値をとる．付録に本モデル全体のアルゴリズムを示す．

ここで注意したいのは，本稿では主体の行動をモデル化するために強化学習を用いたが，価値は学習されるというよりは，むしろ内生的に形成されるものであるということである．主体は，より価値の高い財を受け取るという方針に従って交換を行う．ここである主体にとって価値の高い財とは，消費財を除けば彼の消費財の生産者が受け取ってくれるような財である．すなわち，ある財の価値は，他主体がどのような価値関数を用いるかに依存する．他主体も同様に，自分以外の他主体がどのような価値関数を用いるかに依存して，自分の価値関数をとときどき刻々修正する．このため，ある財の実際の価値（もしそう呼ぶならば）は，各主体が各財に与える価値の推定値と切り離すことができない．

## 6. シミュレーション

本章では，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．パラメータは、各々  $\alpha = 0.005$ ,  $R = 1.0$ ,  $N = M = 20$ ,  $TN = 100$ ,  $PC = 0.1$  に設定した．各財の価値については、初期においては未知であるので、 $0.0 \sim R$  の区間の実数をランダムに割り当てる．ただし前章で述べたように、消費財と生産財については、それらの価値を各々  $R$  と  $PC$  に固定した．また各主体の所有財については、初期状態では生産財を所有しているものとする．以上の初期設定のもとでシミュレーションを行った．なお割引率  $\gamma$  に関しては、その値を様々に変えて試行した．

### 6.1 貨幣の生成

本節では、割引率を  $\gamma = 0.33$  に設定した場合のシミュレーション結果を示す．ここで財の貨幣性を表す尺度として大域性という概念を導入し、ある日におけるある財の大域性を、その日に交換過程を実行した主体総数  $N \times TN$  人のうちで、その財を所有あるいは需要した主体の割合と定義する．図 2 は、ある試行における各財の大域性の時間発展を表す．また図 3 は、最も高い大域性を持つ財の財名の時間発展を示す．なおこの試行において、シミュレーションは数十万日にわたって行ったが、図 2 と図 3 では 10,000 日までの結果のみを表示している．また図 2 の縦軸に関して、大域性は  $0.0 \sim 1.0$  の値をとりうるが、見やすさのため  $0.1$  以上の範囲のみを表示する．これらの図から分かるように、シミュレーション開始直後には各財の大域性は低い値を示す．しかし、すぐに財 8 の大域性が増大し  $0.4$  程度の値を示すようになる．この時期は、財 3 の大域性も比較的高い値を示すため、最大大域性を持つ財は目まぐるしく変化する．しかし 7,000 日目あたりから財 8 の大域性が急激に増大し、最高の大域性を持つようになる．この時期以降、財 8 とそれ以外の財の大域性の間に大きな隔たりが生じ、システムの状態は安定する．

上述のように、システムが安定するまでの過渡期において、最高の大域性を持つ財（便宜上貨幣と呼ぶ）の財名は目まぐるしく変化する．図 4 は、この期間における貨幣の寿命分布を示したものである．横軸はある財が貨幣に君臨してから他の財にその地位を譲るまでの在位期間を、また縦軸はそのような在位期間がどのくらいの頻度で観察されるかを示している．両軸は

この試行では、たまたま財 8 が大域性を獲得したがそれは偶然であり、シミュレーションで用いるランダムシードに依存して様々な財が大域的となりうる．またこの試行では、数十万日後においても財 8 が大域的財のまま安定していたが、試行によっては、大域的財の生成と崩壊が延々と繰り返される場合もあることを付記しておく．

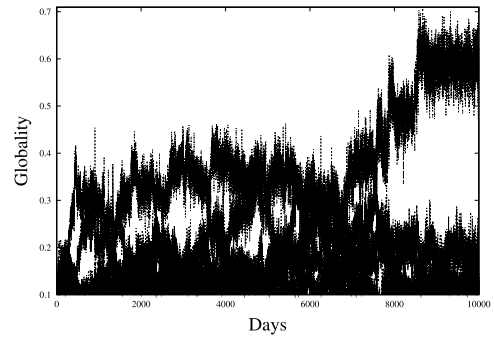


図 2 財の大域性の時間発展

Fig. 2 Time evolution of globality of each good.

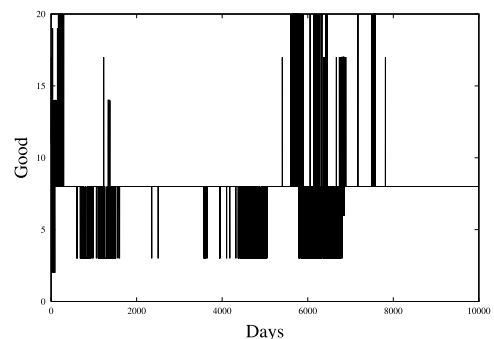


図 3 最高大域性を持つ財名の時間発展

Fig. 3 Time evolution of trade name of the most global good.

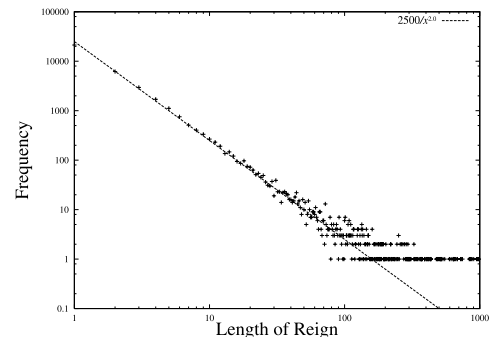


図 4 貨幣の在位期間の頻度

Fig. 4 Log-log plot of distribution of frequency of reign of money.

ともに  $\log$  スケールで表示されている．なおこの結果を得るために、10,000 日までのシミュレーションをランダムシードを変えて 30 回試行した．図から分かるように、貨幣の寿命分布は、傾き  $-2.0$  のべき分布を示す．このような分布は、様々な動物の学習過程や適



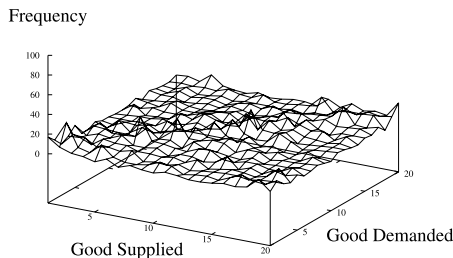


図 5 需給状況 (23 日目):  $\gamma = 0.33$

Fig. 5 Situation of demand and supply (the 23th days:  $\gamma = 0.33$ ).

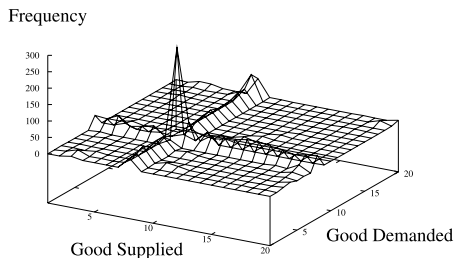


図 7 需給状況 (191,567 日目):  $\gamma = 0.33$

Fig. 7 Situation of demand and supply (the 191,567th days:  $\gamma = 0.33$ ).

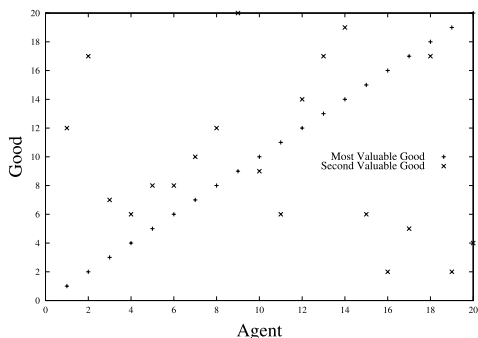


図 6 主体の財に対する価値評価 (23 日目):  $\gamma = 0.33$

Fig. 6 Evaluation for goods (the 23th days:  $\gamma = 0.33$ ).

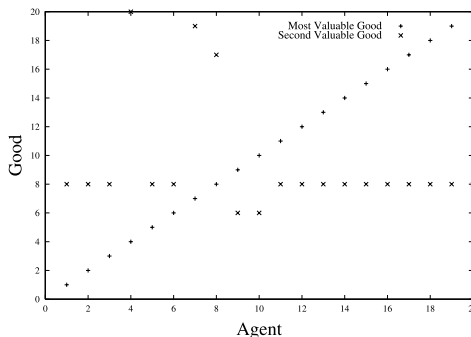


図 8 主体の財に対する価値評価 (191,567 日目):  $\gamma = 0.33$

Fig. 8 Evaluation for goods (the 191,567th days:  $\gamma = 0.33$ ).

応過程において、しばしば観察される<sup>7),11),17)</sup>、。

図 5 は、シミュレーション開始直後のある日 (23 日目) の経済の需給状況を示したものである。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。また 図 6 は、この日に各主体が最も価値が高いと見なした財と 2 番目に価値が高いと見なした財を示す。当然のことながら、すべての主体は消費財に最も高い価値を見出すが、2 番目に価値が高いと見なす財は主体ごとに異なっている。これらの図から、この日は各主体がバラバラの価値関数に従って、様々な財を無秩序に需要していることが分かる。

一方 図 7 は、システムが安定していた時期のある日 (191,567 日目) の需給状況を示したものである。この日、財 8 の大域性は 0.85 という非常に高い値を示す。図から分かるように、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

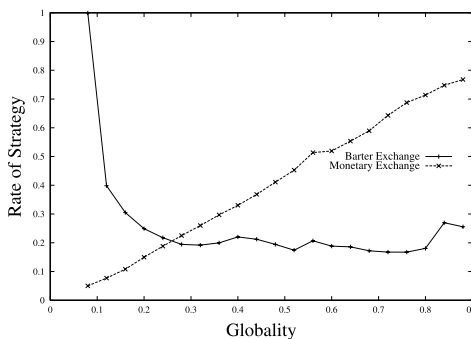


図 9 貨幣の大域性と主体の使用戦略

Fig. 9 Globality of money and strategies of agents.

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。図 8 は、図 6 と同様、この日に各主体がどのような財に価値を見出したのかを示す。ここで注目すべきは、ほとんどの主体が一致して、財 8 を 2 番目に価値が高いと見なしているという点である。

図 9 の横軸は貨幣の大域性を、縦軸はそのような大域性を持つ貨幣が経済内に存在する日に、各主体がどのような交換戦略をとっていたかを表す。ここで、物々交換戦略と貨幣交換戦略という 2 つの戦略を考え

このようなべき分布は、生物種の寿命分布においても見出され、そのような特性を有する進化パターンは断続平衡といわれる<sup>1),16)</sup>。

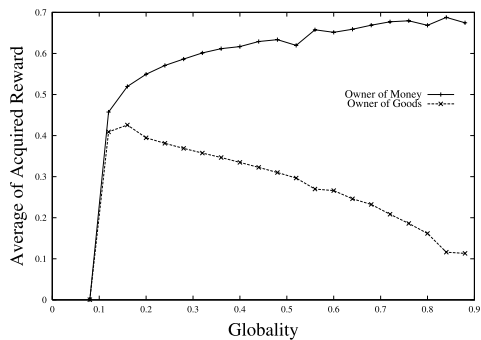


図 10 貨幣の大域性と主体の獲得報酬

Fig. 10 Globality of money and reward acquired by agents.

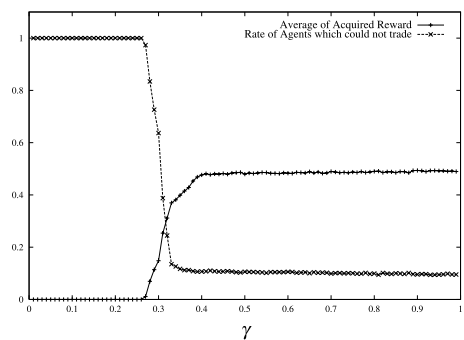


図 11 割引率と経済の交換効率

Fig. 11 Discount rate and exchange efficiency of the trade system.

る。物々交換戦略とは「主体が自分の生産財を所有しているとき、それを消費財と交換しようとする」というものである。一方貨幣交換戦略とは「主体が自分の生産財を所有しているとき、それを貨幣と交換しようとする」というものである。図から分かるように、貨幣の大域性が増加するのにもなって、より多くの主体が貨幣交換戦略を採用するようになる。なおこの図の縦軸の値は、50,000日のシミュレーションをランダムシードを変えて30回試行した結果から得られたものである。この点に関しては以下で示す図10についても同様である。図10は、貨幣の大域性と主体の獲得報酬の関係を表す。この図で横軸は貨幣の大域性を、縦軸はそのような大域性を持つ貨幣が経済内に存在する日に、各主体が獲得した報酬の平均値を示す。なお図では、貨幣の所有者とそれ以外の財（商品と呼ぶ）の所有者の獲得報酬を別々に表示している。この図から、貨幣の所有者は、貨幣が大域的になるに従って、より多くの報酬を得るようになることが分かる。一方商品の所有者の獲得報酬は、大域性の増大にもなって減少していく。以上の結果から、貨幣が大域的になるにつれて、各主体はいったん生産財を貨幣と交換し、その後貨幣と消費財を交換する、つまり貨幣を媒介とした間接交換を行うことで、効率良く消費財を獲得できるようになることが分かる。

6.2 割引率に依存した主体行動の変化

前節では、割引率の値を  $\gamma = 0.33$  に固定してシミュレーションを行った。本節では、割引率の変化にもなって主体の行動がどのように変化するかについて分析する。図11は、割引率  $\gamma$  の値と経済の交換効率の関係を示す。ここでは、経済全体の交換効率を表すために、シミュレーション期間中に各主体が獲得した報酬の平均値と、同じくシミュレーション期間中に交換を行うことができなかった主体の割合という2つの

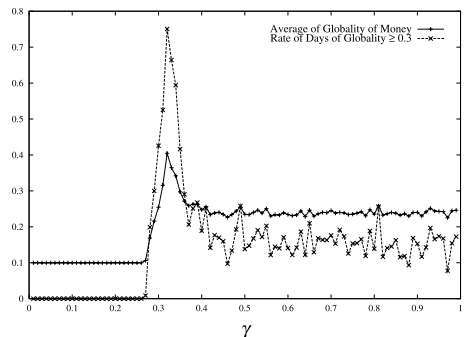


図 12 割引率と貨幣の大域性

Fig. 12 Discount rate and globality of money.

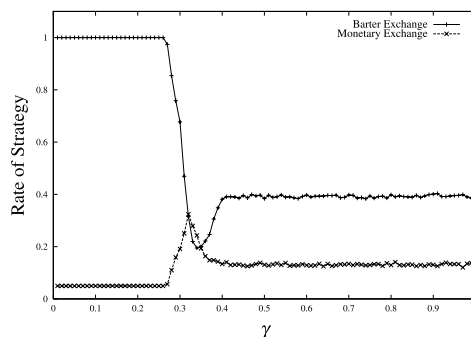


図 13 割引率と主体の使用戦略

Fig. 13 Discount rate and strategies of agents.

尺度を用いる。ここで、交換を行うことができなかった主体数は、交換過程を実行した主体総数から、交換できた主体数と交換しなかった、すなわち所有財を貯蓄した主体数を差し引いたもので表される。なおこの図で横軸の  $\gamma$  の値は、0.01 から 0.01 刻みで 0.99 まで変化させた。一方縦軸の値は、各  $\gamma$  の値に対して10,000日のシミュレーションをランダムシードを変えて30回試行したものを平均したものである。この点に関しては、以下で示される図12と図13についても同様である。図11に示されたように、割引率の

値が小さいとき、主体はまったく交換を行うことができず、獲得報酬は0のままである。しかし  $\gamma = 0.27$  のあたりから次第に報酬を獲得し始め、 $\gamma = 0.4$  まで獲得報酬は上昇し続ける。しかしそれ以降上昇は止み、一定値に落ち着く。この領域においては、ほとんど交換を拒否されることはない。このように、割引率のパラメータ領域は、まったく交換が生じない領域 ( $\gamma \leq 0.26$ ) と、非常に高い交換効率が達成される領域 ( $\gamma \geq 0.4$ ) と、両者の中間領域 ( $0.26 < \gamma < 0.4$ ) に分けられる。

図 12 は、割引率の変化にともなって経済内に存在する貨幣の大域性がどのように変化するかを示したものである。ここではシミュレーション期間を通しての貨幣の大域性を表す尺度として、各日の貨幣の大域性の平均値と、シミュレーション期間中で貨幣の大域性が 0.3 を超えた日の割合を用いる。どちらの尺度も中間領域で最も高い値を示す。また 図 13 は、割引率に依存して主体の交換戦略がどのように変化したかを示す。貨幣の大域性と同様に、貨幣交換戦略を用いる主体の割合は、中間領域で最も高くなる。一方物々交換戦略を用いる主体の割合は、貨幣交換戦略の場合とは逆に、割引率の増加にともなっていったん減少するものの、再び上昇に転じ  $\gamma \geq 0.4$  の領域で一定値をとる。これら結果から、貨幣は中間領域において最も生成されやすいと結論できる。しかしながら、 $\gamma \geq 0.4$  の領域では貨幣が生成されにくいにもかかわらず、なぜ高い交換効率が実現されているのだろうか。以下で各領域におけるシステムの状態を詳しく見ていこう。

まず 図 14 は、割引率を  $\gamma = 0.2$  に設定してシミュレーションを行ったとき、ある主体が各財に認めた価値の時間発展を示す。他の主体の場合も同様の振舞いを示すので、ここでは 1 人の主体の価値評価のみを表示する。図から分かるように、各財の価値は、1.0 の価値を持つ消費財を除いて、すべての財が 0.1 程度かそれ以下の一定値をとる。つまり、主体は自分の消費財のみに高い価値を見出し、それ以外の財に対しては、生産財と同等かそれ以下の価値しか認めない。このため、図 15 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ことになる。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。

次に 図 16 は、 $\gamma = 0.33$  に設定したときに、ある

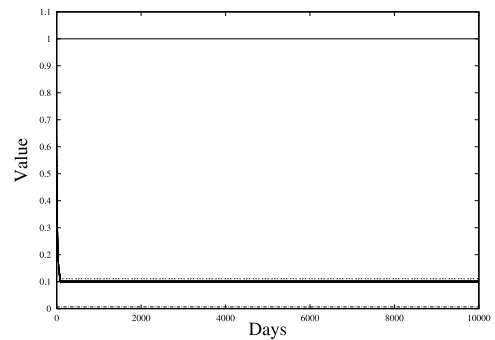


図 14 ある主体が各財に認めた価値:  $\gamma = 0.2$

Fig. 14 Value of each good appreciated by an agent:  $\gamma = 0.2$ .

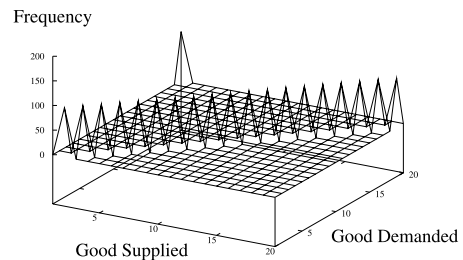


図 15 需給状況:  $\gamma = 0.2$

Fig. 15 Situation of demand and supply:  $\gamma = 0.2$ .

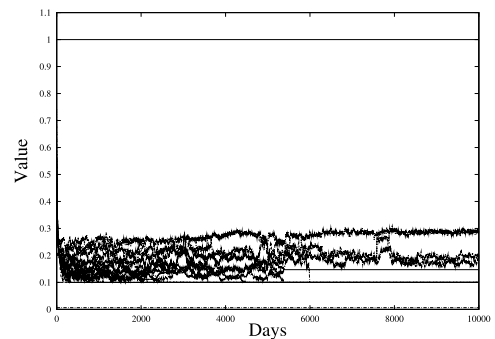


図 16 ある主体が各財に認めた価値:  $\gamma = 0.33$

Fig. 16 Value of each good appreciated by an agent:  $\gamma = 0.33$ .

主体が各財に認めた価値の時間発展を表す。なおこの結果は、図 2 で示したシミュレーションから得られたものである。この試行においては、7,000 日目あたり以降、財 8 が貨幣として使用されるようになる。図 16 に示したように、この主体は学習が進行するにつれて、消費財に最も高い価値を、財 8 に 2 番目に高い価値を見出すようになる。他の多くの主体に関してもこの傾向は変わらない。この時期、経済の需給状況は、図 7 に示したように財 8 を中心とした十字の構造を示す。

図 13 から明らかなように、この点は、シミュレーションで用いるランダムシードに依存しない。

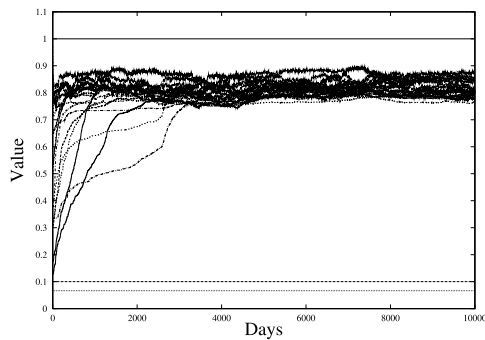


図 17 ある主体が各財に認めた価値： $\gamma = 0.9$

Fig. 17 Value of each good appreciated by an agent:  
 $\gamma = 0.9$ .

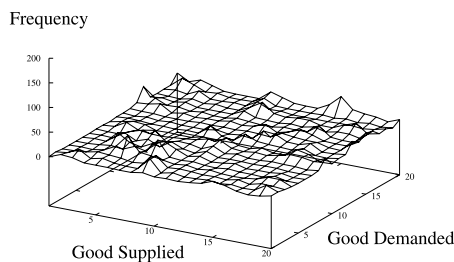


図 18 需給状況： $\gamma = 0.9$

Fig. 18 Situation of demand and supply:  $\gamma = 0.9$ .

つまりこの領域においては貨幣が生成する。

次に図 17 と図 18 は、 $\gamma = 0.9$  に設定したときのシミュレーション結果を表す。図 18 は、シミュレーション期間中で各主体の平均獲得報酬が最も高かった日の需給状況を示す。この日は、非常に交換効率が低いにもかかわらず、十字の構造、すなわち貨幣は見られない。図 17 に示されたように、主体は生産財以外のほとんどの財に、0.8 程度の非常に高い価値を認めている。これは、主体が生産財を所有している場合、相手が何を所有していても交換に応じるということを意味する。このため、各主体は交換を拒否されることがほとんどなく、結果的に多くの報酬を獲得することができる。

## 7. 議論とまとめ

本稿では、強化学習主体による貨幣の生成過程を論じた。本研究によって、単純な学習能力しか持たない

主体間の相互作用によっても貨幣が生成することが明らかになった。本モデルにおける強化学習主体は、Menger<sup>10)</sup> によっても指摘されたように、経済内に販売力の高い財が存在するならば、たとえそれが消費財でなくても受け取るようになる。つまり主体は、消費財を直接需要するのではなく、最も販売力に富む財を媒介とした間接交換を行うことで、効率良く消費財を獲得するようになる。ただし、本モデルにおいて必ず貨幣が生成するわけではない。貨幣が生成するか否かは主体の学習能力に依存する。5 章でも述べたように、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、主体は長期的観点から、将来の報酬を現在の報酬と同等に評価する。

割引率の値が小さいとき、主体は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する。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、各主体の獲得報酬はすべて 0 となる。

一方、割引率の値が大きいとき、主体は長期的観点から、消費財以外の財にも高い価値を見出す。生産財は、自ら労働力を投入して生産したものである。また消費財は、それを消費することによって報酬が得られる。このため、これら 2 つの財は主体にとって実質的価値を持つといえる。しかしながらそれ以外の財は、主体にとって元来無価値であるはずである。そうであるにもかかわらず、主体はそれらの財と自らの生産財を交換する。つまり各主体は、他主体の要求に応じて無償で生産財を提供する。ここでは、他主体に生産財を与える代わりに、他主体から消費財を与えられるという互酬的ネットワークが成立し、各主体は非常に高い効率で消費財を入手することができる。ただし、このような互恵社会を維持していくためには、各主体が将来の報酬を現在の報酬と同等に評価する必要がある。しかし、たとえば食糧や水など、与えられるまで気長に待つことなどできない財もあるだろう。また遠い未来のことは不確実であり、天候不順や戦争などの外的要因によって、報酬を獲得できなくなるかもしれない。このような場合、主体が短期的な損得に関心を

この領域においては、シミュレーションで用いるランダムシードに依存して、財 8 以外にも様々な財が貨幣となりうる。しかし貨幣が生成する場合には、経済の需給状況に十字の構造が見られ、各主体が貨幣に 2 番目に高い価値を見出すという傾向は変わらない。

これらの結果に関しても、シミュレーションで用いるランダムシードには依存せず、一般的にこのような傾向が見られる。

ここで注意したいのは、このような互恵社会においても、主体が利他的であるわけではないということである。物々交換、貨幣交換、互恵のいずれの社会においても、主体は自らの価値関数に従って、自分の所有財をより価値の高い財と交換するというグリーディ方針に従っているにすぎない。ここで述べた社会形態の違いは、システム外部の我々がシステムに与えた 1 つの解釈である。

持つようになることは避けられない。このため現実世界において互惠社会を維持していくためには、西部<sup>12)</sup>が指摘するように、ある種の規範やサンクションが必要なかもしれない。

貨幣は、物々交換社会と互惠社会の中間領域において出現する。この領域において、各主体は物々交換社会のように消費財以外の財をまったく受け入れないわけでも、また互惠社会のようにすべての財を受け入れるわけでもなく、自分の生産財をある特定の財とのみ交換する。このような貨幣交換システムを創発することによって、社会は、交換社会でありながら、互惠社会に近い交換効率を実現しうる。これは、4章で示したように、物々交換社会と互惠社会のギャップは、貨幣の導入によって無効化される、という分析結果とも整合的である。

1章でも述べたように、これまで欲望の二重の一致の困難とその解消という枠組みのもとで、物々交換社会と貨幣交換社会の関係については数多く論じられてきた。また互惠社会と貨幣交換の関係について論じた研究もある<sup>18)</sup>。本研究では、物々交換社会、貨幣交換社会、互惠社会という3つの社会を、割引率という単一のパラメータによって関連付けることができた。つまり割引率の値に依存して、異なる3つの社会形態が出現しうる。ただし、これらの社会の出現に関して、時間的な先後関係をつけることはできない。確かに本モデルのシミュレーションにおいて、主体の学習が進行するにつれて、貨幣が生成することが示された。しかし、貨幣交換は物々交換から出現するわけではない。本モデルにおいてこれらの社会形態は、各パラメータ領域において独立に出現する。これらの形態を時間軸上に位置付けることに関しては、今後の課題としたい。また本モデルで貨幣交換システムは中間領域において出現するが、その範囲は狭い。今後は、生産コストPCの変化にともなって、貨幣の出現範囲がどのように拡大あるいは縮小するのかということについても分析を進めていく予定である。

謝辞 本研究は日本学術振興会特別研究員奨励費の研究助成を受けて行われたものであり、ここに謝意を表します。

## 参 考 文 献

- 1) Bak, P. and Sneppen, K.: Punctuated Equilibrium and Criticality in a Simple Model of Evolution, *Physical Review Letters*, Vol.71, p.4843 (1993).
- 2) Bower, G. and Hilgard, E.: *Theories of learn-*

*ing*, 5th edition, Prentice-Hall Inc., Englewood Cliffs, New Jersey (1981). 梅本亮夫 (監訳): 学習の理論—原書第5版, 培風館 (1988).

- 3) Brown, P.M.: Experimental evidence on money as a medium of exchange, *Journal of Economic Dynamics and Control*, Vol.20, pp.583–600 (1996).
- 4) Duffy, J. and Ochs, J.: Emergence of money as a medium of exchange: an experimental study, *American Economic Review*, Vol.89, No.4, pp.847–877 (1999).
- 5) Iwasaki, A., Imura, S., Oda, S., Hatono, I. and Ueda, K.: Does Reinforcement Learning Simulate Threshold Public Goods Games?: A Comparison with Subject Experiments, *IEICE Trans. Inf. & Syst.*, Vol.E86-D, No.8, pp.1335–1343 (2003).
- 6) Jones, R.A.: The Origin and Development of Media of Exchange, *Journal of Political Economy*, Vol.84, No.4, pp.757–775 (1976).
- 7) Kitabayashi, N., Kusunoki, Y. and Gunji, P.Y.: The emergence of the concept of a tool in food-retrieving behavior of the ants *Formica japonica* Motschulsky, *BioSystems*, Vol.50, pp.143–156 (1999).
- 8) Kiyotaki, N. and Wright, R.: On money as a medium of exchange, *Journal of Political Economy*, Vol.97, No.4, pp.927–954 (1989).
- 9) Marimon, R., McGrattan, E. and Sargent, T.J.: Money as a medium of exchange in an economy with artificially intelligent agents, *Journal of Economic Dynamics and Control*, Vol.14, pp.329–373 (1990).
- 10) Menger, C.: *Grundsätze der Volkswirtschaftslehre*, Wilhelm Braumüller, Wien (1871). 安井琢磨, 八木紀一郎 (訳): 国民経済学原理, 日本経済評論社 (1999).
- 11) Mizukami, E., Gunji, P.Y. and Migita, M.: Learning Process by goldfish and its use of a local site as a map, *BioSystems*, Vol.54, pp.91–104 (1999).
- 12) 西部 忠: 互酬的交換と等価交換—再生産経済体系における価格の必要性, 経済学研究 (北海道大学), Vol.47, No.1, pp.25–42 (1997).
- 13) 西部 忠: 自律分散型市場における貨幣, 経済学研究 (北海道大学), Vol.50, No.3, pp.78–97 (2000).
- 14) 大内 東, 山本雅人, 川村秀憲: マルチエージェントシステムの基礎と応用—複雑系工学の計算パラダイム, コロナ社 (2002).
- 15) Pearce, J.M.: *An introduction to animal cognition*, Lawrence Erlbaum Associates Ltd. (1987). 石田雅人ほか (訳): 動物の認知学習心理学, 北大路書房 (1990).

- 16) Raup, D.: *Extinction: Bad Genes or Bad Luck?*, W.N. Norton & Company (1991).
- 17) Shimada, I., Minesaki, Y. and Hara, H.: Temporal fractal in the feeding behavior of *Drosophila melanogaster*, *Journal of Ethology*, Vol.13, pp.153–158 (1995).
- 18) Shinohara, S. and Gunji, P.Y.: Emergence and collapse of money through reciprocity, *Applied Mathematics and Computation*, Vol.117, pp.131–150 (2001).
- 19) Sutton, R.S. and Barto, A.G.: *Reinforcement Learning: An introduction*, The MIT Press, Cambridge, Massachusetts (1998). 三上貞芳, 皆川雅章 (訳): 強化学習, 森北出版 (2000).
- 20) 安富 歩: 貨幣の複雑性生成と崩壊の論理, 創文社 (2000).
- 21) 吉沢英成: 貨幣と象徴, ちくま学芸文庫 (1993).

## 付録 モデルのアルゴリズム

```

初期化
すべての主体  $i \in A$  に対してその所有財を  $\eta_i = \pi_i$  とする
すべての主体  $i \in A$  とすべての財  $\eta \in G$  に対して
価値関数  $V_i(\eta)$  の値を初期化
各ターンに対して繰り返し:
  交換過程を実行する順番をランダムに決める
  各主体に対して繰り返し:
    取引相手の探索
    主体  $i$  に取引相手  $o_i$  が存在するとき
      主体  $i$  の価値関数と所有財の更新
      所有財  $\eta_i$  が生産財  $\pi_i$  のとき
         $V_i(\pi_i) \leftarrow PC$ 
         $\eta_i \leftarrow \eta_i$ 
      所有財  $\eta_i$  が生産財  $\pi_i$  でないとき
        獲得財  $\theta_{o_i}$  が消費財  $\theta_i$  でないとき
           $V_i(\eta_i) \leftarrow V_i(\eta_i) + \alpha[\gamma V(\eta_{o_i}) - V_i(\eta_i)]$ 
           $\eta_i \leftarrow \eta_{o_i}$ 
        獲得財  $\theta_{o_i}$  が消費財  $\theta_i$  のとき
           $V_i(\eta_i) \leftarrow V_i(\eta_i) + \alpha[\gamma V(\theta_i) - V_i(\eta_i)]$ 
           $V_i(\theta_i) \leftarrow R$ 
           $\eta_i \leftarrow \pi_i$ 
    取引相手  $o_i$  の価値関数と所有財の更新
    所有財  $\eta_{o_i}$  が生産財  $\pi_{o_i}$  のとき
       $V_{o_i}(\pi_{o_i}) \leftarrow PC$ 
       $\eta_{o_i} \leftarrow \eta_i$ 
    所有財  $\eta_{o_i}$  が生産財  $\pi_{o_i}$  でないとき
      獲得財  $\theta_i$  が消費財  $\theta_{o_i}$  でないとき
         $V_{o_i}(\eta_{o_i}) \leftarrow V_{o_i}(\eta_{o_i}) + \alpha[\gamma V_{o_i}(\eta_i) - V_{o_i}(\eta_{o_i})]$ 
         $\eta_{o_i} \leftarrow \eta_i$ 
      獲得財  $\theta_i$  が消費財  $\theta_{o_i}$  のとき
         $V_{o_i}(\eta_{o_i}) \leftarrow V_{o_i}(\eta_{o_i}) + \alpha[\gamma V_{o_i}(\theta_{o_i}) - V_{o_i}(\eta_{o_i})]$ 
         $V_{o_i}(\theta_{o_i}) \leftarrow R$ 
         $\eta_{o_i} \leftarrow \pi_{o_i}$ 
  主体  $i$  に取引相手  $o_i$  が存在しないとき
    主体  $i$  の価値関数の更新
    所有財  $\eta_i$  が生産財  $\pi_i$  のとき
       $V_i(\pi_i) \leftarrow PC$ 
       $\eta_i \leftarrow \eta_i$ 
    所有財  $\eta_i$  が生産財  $\pi_i$  でないとき
       $V_i(\eta_i) \leftarrow V_i(\eta_i) + \alpha[\gamma V_i(\eta_i) - V_i(\eta_i)]$ 
       $\eta_i \leftarrow \eta_i$ 
    
```

(平成 15 年 8 月 20 日受付)

(平成 15 年 9 月 25 日再受付)

(平成 15 年 10 月 7 日採録)



篠原 修二

昭和 42 年生・平成 11 年神戸大学  
大学院自然科学研究科知能科学専攻  
後期博士課程修了・博士(理学)・日  
本学術振興会特別研究員を経て、現  
在株式会社デジタルドリームで Web  
アプリケーション等の開発に従事。