

Title	人工市場アプローチによる介入エージェントを用いた 為替介入効果の分析
Author(s)	松井, 宏樹; 東条, 敏
Citation	人工知能学会論文誌, 20(1): 36-45
Issue Date	2005
Type	Journal Article
Text version	publisher
URL	http://hdl.handle.net/10119/7825
Rights	Copyright (C) 2005 人工知能学会. 松井宏樹, 東条 敏, 人工知能学会論文誌, 20(1), 2005, 36-45.
Description	

人工市場アプローチによる介入エージェントを用いた為替介入効果の分析

Analysis of Foreign Exchange Interventions by Intervention Agent with an Artificial Market Approach

松井 宏樹
Hiroki Matsui

北陸先端科学技術大学院大学情報科学研究科
School of Information Science, Japan Advanced Institute of Science and Technology
hmatui@jaist.ac.jp

東条 敏
Satoshi Tojo

(同 上)
tojo@jaist.ac.jp

keywords: artificial market, multi-agent system, foreign exchange market, reinforcement learning, complex system

Summary

We propose a multi-agent system which learns intervention policies and evaluates the effect of interventions in an artificial foreign exchange market. Izumi et al. had presented a system called AGEDASI TOF to simulate artificial market, together with a support system for the government to decide foreign exchange policies. However, the system needed to fix the amount of governmental intervention prior to the simulation, and was not realistic. In addition, the interventions in the system did not affect supply and demand of currencies; thus we could not discuss the effect of intervention correctly. First, we improve the system so as to make much of the weights of influential factors. Thereafter, we introduce an intervention agent that has the role of the central bank to stabilize the market. We could show that the agent learned the effective intervention policies through the reinforcement learning, and that the exchange rate converged to a certain extent in the expected range. We could also estimate the amount of intervention, showing the efficacy of signaling. In this model, in order to investigate the aliasing of the perception of the intervention agent, we introduced a pseudo-agent who was supposed to be able to observe all the behaviors of dealer agents; with this super-agent, we discussed the adequate granularity for a market state description.

1. はじめに

2003 年, 日本政府は外国為替市場において 1 年間で 20 兆円以上もの米ドル買い介入を行っている [財務 04]. 全体的に見ればこの操作が円高を防いだことは明らかである. しかし, これだけ巨額な介入を行っているにもかかわらず, 常に為替レートが円安の状態に安定していたわけではない. ときに大きな円高への変動すら経験している. では, 介入は市場にどのように影響するのであるのか. また, 有効な介入とはどのようなものであるのか.

本研究で用いる人工市場とは, マルチエージェントモデルとして計算機上に作り出された架空の市場のことである. 市場参加者である複数のエージェントを計算機上の市場で取引させることで市場に起こる様々な現象のシミュレーションを行う. 計算機プログラムであるので, 現実には起こりえない状況でのシミュレーションを行ったり現実では知ることの難しい各市場参加者の投機行動などのデータを得ることが可能である. この特徴を活かして, 従来の経済理論では説明できなかった市場現象の解析などが行われている. また, 人工市場をマルチエージェン

トモデルとして分類すると多数のエージェントが共通の目的を達成するために行動するモデルではなく, 各エージェントがそれぞれの目的のために行動する競合モデルということができる. 本研究の目的は, 外国為替市場における介入の影響とそのメカニズムを人工市場アプローチにより分析することである.

本論文では, 2 章で本研究で用いる人工市場モデルについて, 3 章で現実の中央銀行に相当する介入エージェントについて説明する. 4 章でこのモデルを用いた介入の効果の検証実験について, 5 章では 4 章で報告した実験の検証についてそれぞれ報告する. 最後に 6 章で考察と今後の課題についてまとめる.

2. 人工市場モデル

本研究では和泉らの作成した AGEDASI TOF*¹ [和泉 00] を基にした人工市場モデルを用いている. 本章では AGEDASI TOF とその改良点について述べる.

*1 A Genetic-algorithmic Double Auction Simulation in Tokyo Foreign exchange market

表 1 予想材料

k	予想材料 $x^k(t)$	もともになる生データ
1	景気	[米][日] GDP etc.
2	物価	[米][日] 消費者物価指数 etc.
3	金利	[米][日] 公定歩合, 長期金利
4	マネーサプライ	[米][日] マネーサプライ
5	貿易収支	[米][日] 貿易収支
6	雇用	[米] 失業率 etc.
7	個人消費	[米] 小売売上, 個人所得
8	介入	[米][日] 介入
9	要人発言	[米][日] 中銀総裁の発言 etc.
10	マルク	ドル/マルク, 円/マルク
11	石油	石油価格
12	政治	政情, 国際的な事件
13	株	[米][日] 株価
14	債券	[米][日] 債券価格
15	短期トレンド 1	先週の変動値 ($\Delta R(t-1)$)
16	短期トレンド 2	変動の変動値 ($\Delta R(t-1) - \Delta R(t-2)$)
17	長期トレンド	5 週間の変動 ($R(t-1) - R(t-6)$)

$R(t)$: 期間 t の為替レート (対数)
 $\Delta R(t) = R(t) - R(t-1)$

2.1 AGEDASI TOF

AGEDASI TOF は, 100 人のエージェントからなるマルチエージェントシステムを用いた円・ドルを扱う人工外国為替市場モデルである。各エージェントは円・ドル資産を持ち, その割合をレートに合わせて変化させることで利益を得ようとするディーラーを模している。

§ 1 入力データ

入力されるデータは新聞記事 [日本 02] および相場解説記事 [国際 02] に基づいて和泉らが景気, 物価, 金利, 介入などの 17 項目のレート変動要因の変化を 7 段階 ($\pm 3, \pm 2, \pm 1, 0$) にコーディングしたものである (表 1)。正の値は伝統的な経済理論に従うとドル高要因となる予想材料であり, 負の値はドル安要因となる予想材料を意味している。また, 17 項目の内, 1~14 の項目はファンダメンタルズ要因*2であり過去の相場解説記事などを元にした外部から入力される予想材料である。一方, 15~17 の項目はトレンド要因*3であり, 長期と短期のトレンドを過去のレートから計算しコーディングした予想材料である。

一期間は現実世界の一週間に対応する。レート決定が一回行われる一期間では次の 5 つのステップが実行される (図 1)。

§ 2 各ステップの内容

i. 知覚ステップ

知覚ステップでは各エージェントが為替レートの変動要因である予想材料 (表 1 の) $x^k(t)$ を知覚する。AGEDASI TOF では全市場参加者の知覚は同じであると仮定されているため, 全エージェントが同じデータを受け取る。

*2 為替レートに影響を与える経済の基礎的条件。これらは為替市場の外から入ってくる情報である。

*3 為替レートの動きの方向性のことでチャート分析によって得られる。これらは為替市場内部の情報である。

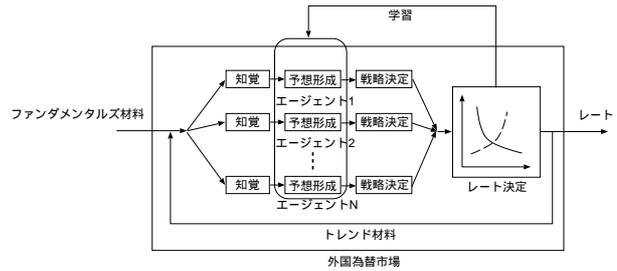


図 1 AGEDASI TOF のフレームワーク

ii. 予想形成ステップ

各エージェントのレートに対する予想形成は, レート (対数*4) $R(t)$ の変動 $\Delta R(t)$ に対する予測値 $E_i[\Delta R(t)]$ を求めることで行われる。各エージェントはそれぞれ独自の市場観を持っており, それは 17 種類の予想材料に対する重みづけで表される。各エージェントの予測レート変動値は, 各予想材料とその予想材料に対する重みづけの積の和として定義される。

$$E_i[\Delta R(t)] \equiv \alpha \left(\sum_{k=1}^{17} x^k(t) w_i^k(t) \right) \quad (1)$$

- $R(t)$: 期間 t の為替レート (対数)
- $E_i[\Delta R(t)]$: エージェント i の $\Delta R(t)$ 予測値
- $x^k(t)$: 期間 t の予想材料 k の値
- $w_i^k(t)$: エージェント i の予想材料 k に対する重みづけ
($\pm 3, \pm 1, \pm 0.5, \pm 0.1, 0$ の 9 段階)
- α : スケール係数

iii. 戦略決定ステップ

戦略決定ステップでは各エージェントが各自の予想に基づき期待収益を最大にする最適ドル資産保有高を計算し, ドルの売買要求量を決定する。

各エージェントの売買戦略は, 期間 t における最適ドル資産保有高が期間 $t-1$ のドル資産保有高よりも大きく (小さく), かつレートが自分の予想したレートよりも安く (高く) 有利な場合には, その差の分だけドルを買って (売って) 保有高を最適値に近づけようとするというものである。

iv. レート決定ステップ

各エージェントの売買戦略を市場全体で集積して, 需要と供給が均衡するような値にモデルの今期のレートを決定する。

v. 学習ステップ

各エージェントの各予想材料に対する重みづけ $w_i^k(t)$ を遺伝子, 重みづけの列 $w_i(t) = (w_i^1(t), \dots, w_i^{17}(t))$ を染

*4 レートの変動を元のレートに対する割合で考えるためにレートを対数でとる。これは, 100 円/ドルから 1 円変動することと 200 円/ドルから 2 円変動することは等価であるという考えに基づく。

色体とみなし、学習ステップでは Simple GA [Goldberg 89] に基づいた遺伝的アルゴリズムを用いて各エージェントは情報交換、学習を行い、予想材料に対する重みづけの列を変化させる。学習ステップで行われる操作は淘汰、交叉、突然変異の3つである。

まず、エージェント i の期間 t における適応度 $*5 F(w_i(t))$ を以下のようにレートの予測と実際のレートのずれをもとに定義する。

$$F(w_i(t)) = -|E_i[\Delta R(t)] - \Delta R(t)| \quad (2)$$

淘汰とは、確率 G (Generation Gap) で選ばれたエージェント i (全エージェント数 $N \times G$ 人) が、適応度 $F(w_j(t))$ に比例する確率で選ばれた相手 j から重みづけの列をコピーする操作である。つまり、各エージェントは自分の予想が正確であれば変更せず、予想が不正確であれば他のエージェントが持つ適応度の高い重みづけの列に入れ替えるのである。これは予測に失敗したディーラーが成功したディーラーと情報交換を行うことによって自らの予想方式を変更することにあたる。

交叉とは、ランダムに全エージェントをペアにし、それぞれのペアの重みづけの列に対し、確率 P_{cross} で一点交叉を行う操作である。これはディーラー間のコミュニケーションによりお互いの意見を交換しあい、その結果予測方針を変更することにあたる。

最後に突然変異とは、エージェント i がある重みづけ w_i^k を、確率 $P_{mutation}$ でランダムに変化させる操作である。これはディーラーがある予想材料に対して、思いつきで新たな価値を設定することにあたる。

§3 AGEDASI TOF によるシミュレーション

AGEDASI TOF によるシミュレーションは、実験対象であるテスト期間とエージェントのトレーニングを行うトレーニング期間から成る(図2)。まず、テスト期間の前の数年間をトレーニング期間とし、現実のレートとレート予想材料のデータを用いてエージェントのトレーニングを行う。トレーニング期間ではレート決定ステップにおいてレート決定を行わず、現実のレートを系のレートとして用いる。トレーニング期間は20回*6繰り返し、現実のレートを用いた学習によってランダムに生成されたエージェントの重みづけはテスト期間直前の市場参加者の状態に最適化される。以下、このトレーニング期間の後、個々のエージェントが持つ重みづけ列 $w_i(t)$ を初期状態と呼ぶ。トレーニング終了後、現実のレート予想材料のデータを用いてテスト期間のシミュレーションを行う。このシミュレーションによって現れたレートの推移をシミュレーションパス (simulation path) と呼ぶ。テスト期間では現実のレートデータは用いず、各エージェントの売買行動によって決定する。

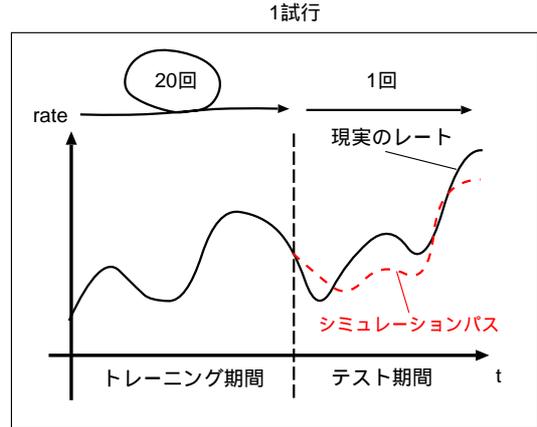


図2 AGEDASI TOF によるシミュレーション

	k	1	2	3	...	16	17	
予想材料	$x^k(t)$	1	-3	0	...	2	0	
agent i	$w_i^k(t)$	3	-0.1	0.5	...	0	-3	適応度: 高
agent j	$w_j^k(t)$	0	1	3	...	1	0.5	適応度: 低

↓

agent i	$w_i^k(t+1)$	3	-0.1	0.5	...	0	-3	
agent j	$w_j^k(t+1)$	3	-0.1	3	...	0	0.5	

図3 提案手法による学習の例

2.2 情報交換の仕組みの改良

人工市場モデルを用いてシミュレーションを行うことのメリットとしてマイクロレベルすなわちエージェントの状態、行動を分析できることがあげられる。しかし、AGEDASI TOF ではエージェントの学習に Simple GA を用いているためにその行動が不自然になってしまっている。まず、交叉において行われる重みづけの列の一点交叉は、対等にレート予想材料の重みづけを入れ替えるという点で現実のディーラーの情報交換としては説明できない不自然な行動である。また、淘汰においてはレートの予想の正確さを高めるために行うにもかかわらず、予想がより正確だったエージェントの重みづけ列をすべてコピーしている。そこで本研究では、AGEDASI TOF における交叉を廃止し、レートの予想に関係した重みづけだけをコピー対象とする淘汰による情報交換の仕組みを導入する。

図3に適応度の低いエージェント j が適応度の高いエージェント i から重みづけをコピーする例を示す。この例で、3つめの予想材料と最後の予想材料は0でありレート予想に無関係であるため、エージェント j はそれらに対する重みづけに関してはコピーを行わず、自分の認識を変えていない。この操作では、適応度に無関係な重みづけをむやみにコピーしないことで各エージェントが個性を失うことを防いでいるとも言える。また、現実のレート予想材料のデータを用いたシミュレーションでは本研究の情報交換の仕組みを用いたモデルの方が AGEDASI TOF よりも精度が高いという結果も得られている [Matsui 03b]。

*5 厳密には、エージェント i の期間 t にとった予想方式(予想材料に対する重みづけの列 $w_i(t)$) の適応度

*6 現実のレートをシミュレートする実験において、十分な精度が得られることから20回とした。

3. 介入エージェントの導入

本研究では前章で述べた市場モデルに政府、中央銀行に相当する介入エージェントを導入する。介入が為替レートに及ぼす効果については2種類あると言われている[渡辺 94]。1つは介入が外国通貨の需給に影響を与え、それを反映してレートが変化するというものである。これはポートフォリオ・バランス効果と呼ばれている。もう1つは中央銀行の介入から市場参加者が将来の金融政策に関する意図を読み取り、それに基づいて為替レートに関する予想を変化させることでレートが変化するというものである。これはシグナル効果と呼ばれている。近年の介入においてはシグナル効果の重要性が指摘されており、ポートフォリオ・バランス効果はシグナル効果と比べるとかなり小さいという事例も報告されている[Dominguez 93, Ramaswamy 00]。しかし、AGEDASI TOFにおいて介入は予想材料の1つにすぎず市場の需給に影響しないため、これらの介入の効果の検証を行えなかった。また、AGEDASI TOFでは実験を行う前にいつ、どれくらい介入するかをあらかじめ決定しておく必要があるため、市場の状態に応じた介入ができなかった。本研究では介入を行う機関をエージェントとして系に組み込むことで、これらの問題点を解消する。

介入エージェントは以下の点で一般の市場参加者であるディーラーエージェントと異なる。

目的： 為替レートを目標範囲内に安定させることを目的とする。ディーラーエージェントと異なり、取引により利得が得られるかどうかとは無関係である。各ステップ(2・1・2節)での行動：

- 戦略決定ステップでは、介入量を決定する。
 - 一般に、介入量は1ディーラーエージェントの注文量よりも大きい。
 - 介入を行うことを市場に知らせる(シグナリング)場合は、介入エージェントの行動が予想材料の介入 $x^8(t)$ として知覚される。
- 学習ステップにおいて、ディーラーエージェントは2・2節で述べたアルゴリズムで学習を行うが、介入エージェントはそれとは独立に強化学習を用いて有効な介入政策を学習する。

[松井 02]では、既存の介入政策をアルゴリズムとして実装することでその有効性を検証するシステムを作成しているが、本研究では自動的に介入政策を獲得するシステムを作成し、様々な条件下での介入の有効性の検証を可能にする。

3・1 強化学習による介入政策の獲得

介入エージェントは、強化学習によりレートの安定に有効な介入政策を獲得することを目的とする。介入エージェントが行う学習はディーラーエージェントにより構成される市場の状態を知覚し、その状態に応じた行動を

学習するものであり、ディーラーエージェントのものと比べてより市場全体を対象にした学習である。本研究で用いる強化学習手法は Profit Sharing [Grefenstette 88] に基づいたものである。Profit Sharing は強化学習の1種で報酬の得られたエピソード^{*7}上のルールに対する重みづけを強化していく方法である。1つのルールは状態 S とその状態で選択される行動 a の組で表される。

状態 S 、行動 a の定義、およびルール (S, a) に対する重みづけ $W(S, a)$ を変化させる方法を以下に示す。

§1 状態

状態 S を、現実の中央銀行が介入の際に考慮する市場状態として以下のように定義する。現在のレートやその予測を考慮するのは当然であるが介入の効果は市場参加者の介入に対する認識に左右されうる[渡辺 94]ため、介入に対する重みづけの平均と標準偏差を加えて状態表現とした。

$$S \equiv (\bar{E}, \sigma_E, \bar{w}^8, \sigma_{w^8}, x^{15}, R) \quad (3)$$

$$\bar{E} = \sum_{k=1, k \neq 8}^{17} \bar{w}^k x^k$$

ここで x^k は k 番目の予想材料、 \bar{w}^k は予想材料 x^k に対する全ディーラーエージェントの重みづけの平均値である。よって S の第1項 \bar{E} はほぼレートの変動に対する予測の平均となる^{*8}。第2項 σ_E は E の標準偏差である。つまりどれだけディーラーの予測が集中しているかを表している。 \bar{w}^8, σ_{w^8} は全ディーラーエージェントの介入に対する重みづけの平均と標準偏差である。また、 x^{15} は短期トレンド $\Delta R(t-1)$ (表1参照)、 R は現在の為替レートである。計算機実験を行う上で現実的な計算量に収まるようにこれらの値はそれぞれ5~11段階の離散的な値とした。各離散値は、実験過程で各値の変化を観察しその範囲と頻度をもとに設定したものである。

この状態表現を用いることでレートが目標値より高く(低く)かつ上がり(下がり)トレンドならドル売り(買)介入を行う政策(対称的 leaning-against-the-wind 政策[渡辺 94])を学習することも可能になる。

§2 行動

行動 a は「どのように介入するか」を表す。すなわち、介入の向き(ドル買いかドル売るか)とその量である。

$$a \equiv \frac{3q}{Q} \quad (4)$$

a は7段階の離散的な値($\pm 3, \pm 2, \pm 1, 0$)^{*9}をとる。 q が実際の介入量であり、介入エージェントの注文量となる。このとき、注文レートは常に目標範囲の中央値とした(この値を以下、目標レートと呼ぶ)。 Q は最大介入量で、 a

*7 エピソードとは選択されたルールの列である。

*8 介入エージェントが操作できる介入の項目 x^8 が除かれているので全ディーラーの注文レートの平均とは完全には一致しない。

*9 表1の他の予想材料に合わせてこの値とした

が ± 3 をとったときに介入量 q が $\pm Q$ となる．この設定で介入エージェントは「レートが目標レートより低く、 $a > 0$ ならばドルを買う」、「レートが目標レートより高く、 $a < 0$ ならばドル売る」というような介入を行うことができる．

また介入を行うことを市場に知らせる（シグナリング）場合は、行動 a の値がそのまま x^8 としてディーラーエージェントに知覚される．シグナリングを行わない場合は、 x^8 は常に 0 となる．行動 a は各ルールの重みづけ $W(S, a)$ によるルール選択とした．

§3 報酬

各ルールは以下の 3 つの方法で評価され、ルールに対する重みづけ $W(S, a)$ を変化させる．

試行全体の評価

テスト期間全体でレートが目標範囲に入っていたら、エピソード上すべてのルール^{*10}の重みづけを一様に増加させる．

$$W(S, a) \leftarrow W(S, a) + r \quad (5)$$

各週の評価

式 (5) によるルールの重みづけの更新は、1 試行で 1 回しか行われず報酬を得るための条件も厳しいため、式 (5) のみによる重みづけの更新では学習が進みにくいと予想される．そこで本研究では、学習速度の低下を防ぐために各週で実行されたルールを以下の方法で評価し、そのルールの重みづけを随時更新する．

- (1) 各週において決定されたレートが目標範囲内かどうかでその週のルールの重みづけを変化させる．

レートが目標範囲内の場合

$$W(S, a) \leftarrow W(S, a) + \frac{r}{2} \quad (6)$$

レートが目標範囲外の場合

$$W(S, a) \leftarrow W(S, a) - \frac{r}{2} \quad (7)$$

- (2) 介入したにもかかわらず取引できなかった場合、すなわちドル買い介入を行おうとしたが、決定したレートは目標レートよりも高かった、あるいはドル売り介入を行おうとしたが、決定したレートは目標レートよりも安かった場合には、介入行動が市場の需給に全く影響していないだけでなく、「レートを上げるためにドル買い介入を行う」というような介入の基本概念と照らし合わせても無駄な行動であると考えられる．したがって、このような場合はその週のルールの重みづけを減少させる．

$$W(S, a) \leftarrow W(S, a) - \frac{r}{2} \quad (8)$$

本研究では各ルールの初期重みを 1.0、報酬 r を 0.1 とする^{*11}．試行全体の評価、各週の評価での重みづけ変化値の比、および初期重みと報酬 r の比は実験により学習効率のよい値に設定した．

*10 ここでのエピソード長はテスト期間の長さと同じで、評価されるルールの数はテスト期間の週の数と同じになる．

*11 現実の為替市場を前提にすると介入を行わない行動 $a = 0$ を

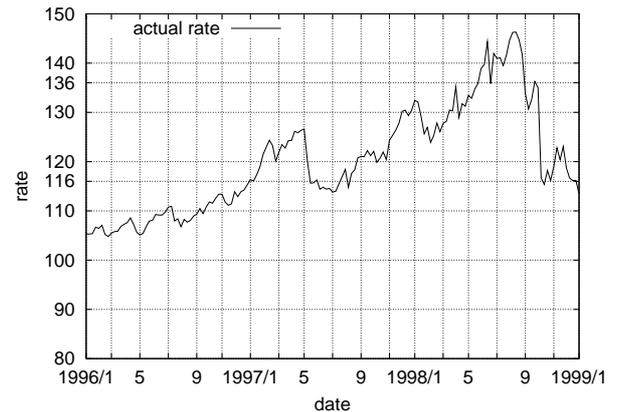


図 4 実験対象期間の現実のレート

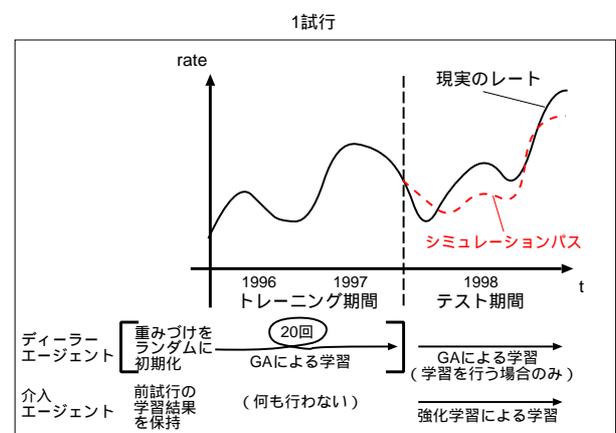


図 5 介入エージェントを用いたシミュレーション

実際の実験では [] 内の過程を保存した初期状態を用いることで省略している．

4. 介入エージェントを用いたシミュレーション

本章では前章で提案した介入エージェントを用いて、有効な介入政策を獲得する実験の報告を行う．また、その有効性の検証のために、

- ランダム介入エージェント ランダムに介入を行う
- 全知覚介入エージェント 全市場参加者の行動を観察し、常に最適な介入を行う
- 全知覚ルールテーブル 全知覚介入エージェントの行動に基づき、強化学習介入エージェントの状態表現の枠組みで作成した介入規則による介入を行う

を順次定義し、本章および 5 章で実験結果を比較する．

4.1 実験の設定

1998 年 10 月、ドル-円レートは 1 週間で 20 円も下落した（図 4）．この 1998 年をテスト期間として実験を行う．まず、1996 年 1 月から 1997 年 12 月の 2 年間で

デフォルトとする設定も考えられるが、本研究では介入の効果を出させることを目的とするため、すべての状態で $a = 0$ を含めた 7 種類の行動からランダムに選ぶ政策を初期政策としている．

表 2 介入量とシグナリング

	最大介入量 Q	シグナリング
$(30\bar{q}_a, \text{yes})$	$30\bar{q}_a$	行う
$(10\bar{q}_a, \text{yes})$	$10\bar{q}_a$	行う
(\bar{q}_a, yes)	\bar{q}_a	行う
$(30\bar{q}_a, \text{no})$	$30\bar{q}_a$	行わない
$(10\bar{q}_a, \text{no})$	$10\bar{q}_a$	行わない
(\bar{q}_a, no)	\bar{q}_a	行わない

\bar{q}_a : 介入エージェントが存在しないシミュレーションでのディーラーエージェントの平均注文量

トレーニング期間として AGEDASI TOF と同様にその期間の現実のレートとレート予想材料のデータを用いてディーラーエージェントのトレーニングを行う。予想材料の重みづけがランダムな状態のディーラーエージェントに対しトレーニング期間を 20 回繰り返すトレーニングを 100 組行い、それぞれを 1998 年直前のディーラーエージェントの初期状態として保存した。その後、テスト期間である 1998 年のシミュレーションを 200,000 試行繰り返した。各試行のテスト期間開始時にはあらかじめ作成した 100 組のディーラーエージェントの初期状態のデータを順に用いた。つまり、100 種類の初期状態についてそれぞれ 2000 回シミュレーションを行うことになる。介入エージェントの各ルールに対する重みづけ $W(S, a)$ の表 (ルールテーブル) は、1 回のシミュレーションごとに初期化は行わず、保持したまま次のシミュレーションを行う。この実験の 1 試行の内容を図 5 に示した。レートの目標範囲 単純化のために介入エージェントのレートの目標範囲は t によらず、常に 116 ~ 136 円とした。

最大介入量 Q とシグナリング 介入量の大小およびシグナリングの有無による介入の効果と比較するために表 2 の設定^{*12}の実験を行った。

ディーラーエージェントの学習 本研究のモデルではテスト期間において、ディーラーエージェントと介入エージェントが相互に学習を行う。このとき、お互いの学習が影響しあうことで有効な戦略を学習することが難しくなる。これは同時学習問題と呼ばれ、マルチエージェント学習の難しさの 1 つとして知られている [荒井 01]。本研究でも、[松井 03a] において同時学習が介入の成功率を下げる原因となることをすでに示している。そこでテスト期間内でディーラーエージェントが学習を行わない場合 (同時学習なし) と行う場合 (同時学習あり) の実験をそれぞれ行った。

評価方法 テスト期間のレートが常に介入エージェントの目標範囲に入っているシミュレーションパス (図

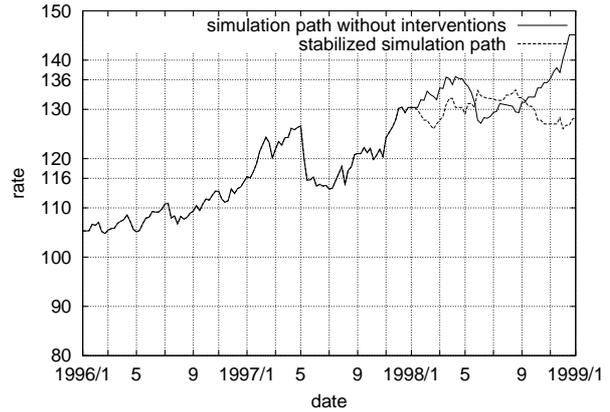


図 6 ディーラーエージェントが学習を行わず介入が行われない場合のある初期状態に対するシミュレーションパスと安定パス

6), すなわち介入が成功したシミュレーションパスを安定パスと呼ぶ。本研究は、シミュレーション回数に占める安定パスの割合 (レートの安定率) で評価を行った。

4.2 実験結果

まず、全体の結果から 1 つの初期状態に対する結果を取り出したものを示す。ディーラーエージェントが学習を行わない場合、介入エージェントが介入を行わなければ確率的な要素が全くないためシミュレーションパスは常に同じになる (図 6)。この初期状態に対するディーラーエージェントが学習を行わない場合の結果を図 7 に、学習を行う場合の結果を図 8 に示す。また、全体の結果を表 3 の「強化学習」欄に示す。全体の結果では、3・1・2 節で定義した 7 種類の介入エージェントの行動をランダムに選択するランダム介入^{*13} (表 3 の「ランダム」欄) と比較した。ここでランダム介入が、注文レートが一定ではないランダムな売買と異なることに注意が必要である。ランダム介入では強化学習を行う介入エージェントと同様に注文レートが目標レートで固定されており、レートが注文レートより高い (安い) 場合、ドル買い (売り) 取引は行わないため介入行動の選択がランダムであってもある程度の効果が期待できる。

以下に実験結果の検証を行う。

介入量

すべての設定で介入量が大いほどレートの安定率が増加している。これは単純に市場内での影響力が増すことが理由であると考えられる。

シグナリング

介入量が同じ場合、ディーラーエージェントが学習を行う場合、行わない場合ともにシグナリングを行った方

*12 1998 年 4 月の東京外国為替市場の一日平均の取引金額は 1,486 億ドル [吉本 00, p. 11] であり、対象期間に近い 1999 年 4 月 10 日に 202 億ドルの介入 [財務 04] が行われたことを考えると、本研究の介入量の設定は妥当な範囲であると考えられる。

*13 介入エージェントの学習を行う前の段階では、7 種類のすべての行動の重みづけが同じであるので、ランダム介入は未学習状態の強化学習介入エージェントによる介入ということもできる。

表 3 全初期状態に対するレート安定率 (100 初期状態 × 100 回平均)

介入エージェントの種類	ディーラーエージェントが学習を行わない場合				ディーラーエージェントが学習を行う場合			
	ランダム	強化学習	全知覚ルールテーブル	全知覚	ランダム	強化学習	全知覚ルールテーブル	全知覚
(30 \bar{q}_a , yes)	53.8 %	69.1 %	75.2 %	100.0 %	48.3 %	57.1 %	73.3 %	99.1 %
(10 \bar{q}_a , yes)	43.7 %	57.5 %	63.0 %	99.0 %	33.3 %	38.8 %	53.4 %	94.0 %
(\bar{q}_a , yes)	35.7 %	47.7 %	50.1 %	97.0 %	19.1 %	23.6 %	31.3 %	85.3 %
(30 \bar{q}_a , no)	53.5 %	56.6 %	58.1 %	94.0 %	45.7 %	48.9 %	59.7 %	95.0 %
(10 \bar{q}_a , no)	40.0 %	42.3 %	43.7 %	80.0 %	27.9 %	28.8 %	35.9 %	83.0 %
(\bar{q}_a , no)	31.2 %	31.5 %	34.5 %	68.0 %	14.3 %	13.6 %	18.9 %	46.9 %

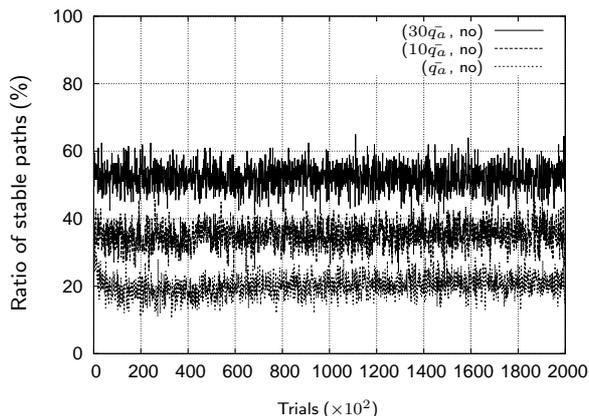
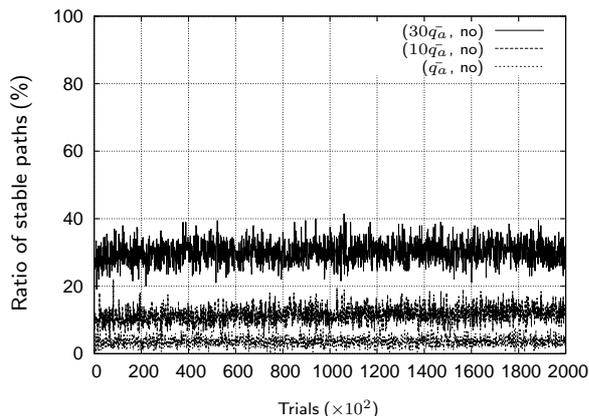
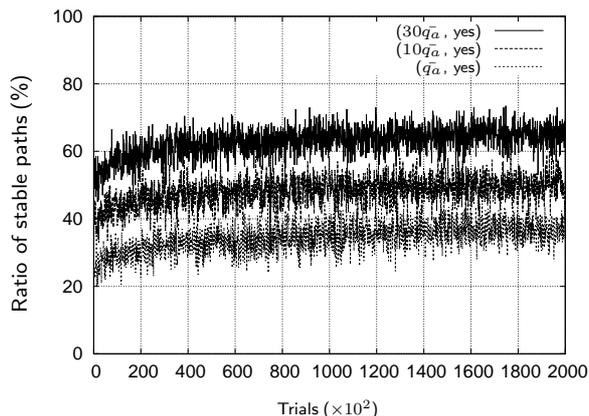
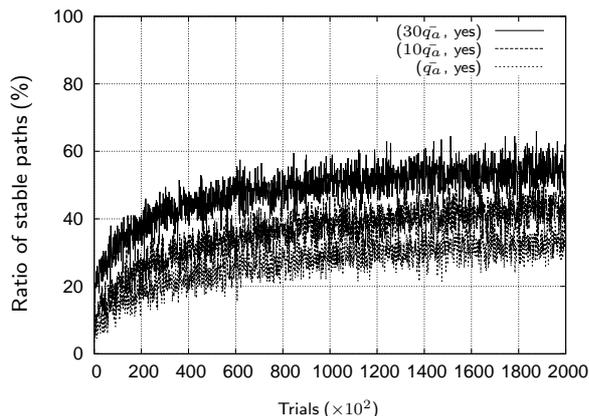


図 7 ある初期状態に対する同時学習が起こらない場合の試行回数とレート安定率 (200 回平均)

図 8 ある初期状態に対する同時学習が起こる場合の試行回数とレート安定率 (200 回平均)

がよりレートを安定させることに成功している．この結果は介入においてシグナル効果の影響が大きいことを示すものである．

同時学習問題

1 初期状態に対する結果で介入量が小さくディーラーエージェントが学習する場合は学習回数が少ないときにレートの安定率が一度、10%ほども低くなっている．これは市場への影響力が小さいために、学習初期に有効な戦略を模索する行動がディーラーの重みづけを攪乱してしまっていると考えられる．この現象はディーラーエージェントが学習する場合のみ起こることであり、同時学習問題が顕著に現れている例である．

また、1 初期状態に対する結果においてこの初期状態では図 6 で示したようにディーラーエージェントが学習を

行わない場合、介入を行わなければ決して安定パスとはならない．一方、同じ初期状態でもディーラーエージェントが学習を行えば確率的な要素があるため、介入を行わなくても安定パスとなる可能性がある．そのために、結果を比較するとディーラーエージェントが学習を行う場合の方が実験開始時、また最終的にもレート安定率が高くなっている．しかし、全体の結果を見るとディーラーエージェントが学習を行う場合は、学習を行わない場合に比べてレートを安定させることができている．これは、同時学習問題に起因していると考えられる．

各エージェントの学習

介入の設定やディーラーエージェントが学習を行うかどうかにかかわらず、介入エージェントが学習で獲得した介入政策では、レートが目標レートよりも低い(高い)

場合、以下の (a) > (b) > (c) > (d) の状態の順によりドル買い (売り) 介入を行う傾向が強かった。

- (a) ・レートが目標レートよりも低く (高く) ,
 - ・下がり (上がり)トレンドで,
 - ・ディーラーエージェントの予測平均がドル安 (高) .
- (b) ・レートが目標レートよりも低く (高く) ,
 - ・ディーラーエージェントの予測平均がドル安 (高) .
- (c) ・レートが目標レートよりも低く (高く) ,
 - ・下がり (上がり)トレンド .
- (d) ・レートが目標レートよりも低い (高い) .

この結果から、介入エージェントは学習結果として対称的 *leaning-against-the-wind* 政策を獲得したことが確認された。また、すぐに変化する可能性のある短期トレンドよりも次週のレートへの影響の大きいディーラーエージェントの予測を重視する介入行動を学習したことがわかる。

次にシグナリングを行う介入エージェントの行動による、ディーラーエージェントの介入に対する重みづけの変化について述べる。本研究の対象期間では、全初期状態中のほとんどのエージェントが介入に対して正の重みづけをしていた。この実験序盤では、介入量が小さい介入エージェントによるほぼランダムな介入に対しては、徐々に負の重みづけを持つディーラーエージェントが増えるという変化が見られた。これは介入の影響力が小さいためにレートがシグナルとは逆の変化をすることが頻繁に起こり、その際にシグナルに従うレートの予測をしたディーラーエージェントが重みづけを変更するためだと考えられる。一方、実験終盤では介入量の設定を問わず重みづけはほとんど変化しなかった。これは介入の効果がシグナル通りに現れているためにディーラーエージェント間で介入に対する重みづけが変更されない状況である。このような市場では介入のシグナル効果が大きく、介入エージェントの行動次第で市場全体をある程度操作することが可能だと考えられる。

5. 介入量と強化学習モデルの検証

本章では、前章で報告した実験における介入量の査定と強化学習モデルの妥当性を検証する。このため、市場の各ディーラーエージェントの行動をすべて第三者的に観察できる介入エージェントを仮想し、各週において最適な介入行動^{*14}をとることができるとしたら、どのくらいレートを安定させることが可能か実験を行った。その後、この理想的な介入を上限と考えたときとの比較において、本実験の設定の実効性と問題を議論する。本章ではこの理想的な介入を行うエージェントを全知覚介入エー

ジェントと名づけ、強化学習介入エージェントと同様に 3.1.2 節で定義した 7 種類の行動を行うものとする。

5.1 介入量の検証

前章の強化学習介入エージェントの代わりに全知覚介入エージェントを用いた実験を行い、各最大介入量の妥当性を検証した。結果を表 3 の「全知覚」欄に示す。

どの設定でも強化学習によって獲得された介入政策より圧倒的に安定率が高かった。このことは現在の介入量の設定でレートをより安定させる介入政策を実行可能であることを示している。

5.2 全知覚ルールテーブルによる介入

次に、全知覚介入エージェントを用いて本研究の強化学習モデル (状態表現、ルールの重みづけの更新方法) の検証を行う。

まず検証の準備として、各介入設定 (介入量、シグナリングの有無) についてそれぞれ 30000 試行 (100 初期状態 × 300 回) の実験を行った。その際に全知覚介入エージェントがとる行動に基づき、強化学習介入エージェントの状態表現の枠組みでルールテーブルを作成する (全知覚ルールテーブルと呼ぶ) この全知覚ルールテーブルは、強化学習モデルの状態表現を用いた際の最適介入政策であると考えられる。全知覚ルールテーブルに基づいて介入することでその際のレートの安定率が高く、5.1 節の全知覚介入エージェントの安定率に近ければ本モデルの状態表現、すなわち介入エージェントの知覚は十分であり、よりレートを安定させるためには重みづけの更新方法などを改良すべきであることがわかる。逆に安定率が低ければ状態表現が不十分であることを示している。また、強化学習介入エージェントによる安定率が、全知覚ルールテーブルによる安定率に近ければ状態表現を除く本強化学習モデルは十分機能していると言える。

実験結果を表 3 の「全知覚ルールテーブル」欄に示す。レート安定率は、すべての設定で全知覚介入エージェントによるものよりも 25 ~ 50 % も低く、現在の状態表現では明らかに不十分であり不完全知覚が問題になっていることを示している。一方で、強化学習介入エージェントの結果は全知覚ルールテーブルによるものには及んでいない。ディーラーエージェントが学習を行わない場合にはその差は最大でも 6 % ほどで、ある程度学習できていると言えるがディーラーエージェントが学習を行い同時学習が起こる場合では最大で 16 % もの差があり、この結果は同時学習問題が起こっていることを実験的に示している。

5.3 状態表現の粒度に関する考察

本節では強化学習モデルにおける状態表現の粒度について議論する。全知覚介入エージェントは、遭遇した状

*14 ここでの最適な介入行動とは、各週のレートを目標レートに最も近づけることができる介入行動である。あくまでも各週のレートのみを考慮する行動であり、長期的なレートの変化は考慮していない。

態すべてに対して最適な行動をとることができた。これはつまり本研究の状態表現と比較すると、より詳細な状態表現によって常に一意に行動を決定することが可能だということである。逆に本研究の状態表現は全知覚介入エージェントのものと比べて非常に粗いために全知覚介入エージェントによってルールテーブルを作成しても、一状態に対する行動を一意に決定することができない。そこで作成した全知覚ルールテーブルを分析し、各状態において最も強化された行動が全知覚ルールテーブルに基づいて行動を選択した場合、どれだけの確率でとられるかを調べた。この確率が 100% に近い状態が多ければ全知覚介入エージェントと同様に行動をほぼ一意に決定することができ、現在の状態表現でも十分であるということができる。

本研究の実験における行動選択の統計をとった結果、どの介入設定でも常に最も強化された行動が選択される状態は全状態数の中の 1.5~10% であり、最も強化された行動が 75% 以上の確率で選択される状態も全状態数の中の 20~35% と低かった。これでは、各状態において行動の選択にばらつきが生じることになり常に最適な介入を行うことが難しい。このことから、やはり介入エージェントの現在の状態表現では知覚が不十分であることがわかる。しかし、本研究の状態表現は現実の市場に則したものであるため、これは現実の中央銀行の知覚能力に限界があることを適切に表現するものである。

6. おわりに

本研究では、現実の政府・中央銀行に当たる介入エージェントを人工外国為替市場モデルに導入することで介入効果の分析、また強化学習により自動的に有効な介入政策を獲得するシステムを作成し、このシステムを用いて為替介入におけるシグナル効果を検証した。その結果、ディーラーエージェントが介入を知覚しない場合に比べて知覚する場合の介入の効果は大きく、介入においてシグナル効果の影響は大きなものであることを示した。

本研究で提案した強化学習介入エージェントによる介入は、介入量が十分であるにもかかわらず全知覚介入エージェントと比較して安定率が大きく下回っている。これは 5.2 節および 5.3 節で検証したように本研究で定義した強化学習の状態表現、すなわち現実の中央銀行の知覚能力の限界を示している。また、ディーラーエージェントが学習を行う場合には同時学習問題が起こり、有効な政策の学習が困難になるため強化学習介入エージェントによる介入の安定率は全知覚介入テーブルによる介入 (5.2 節) よりも低くなっている。特に介入量が少なくシグナリングを行わない場合は、顕著でランダム介入よりも低くなっている。

これらの問題はマルチエージェント学習の課題であり、解決が困難であると考えられる。しかし、本研究の強化

学習介入エージェントはシグナリングを行う場合には行わない場合に比べてレートを安定させることに成功している。これは、大きな力を持つ介入エージェントがうまく他のエージェントの行動を誘導しているからだと思われる。本研究の対象である外国為替市場における介入では介入エージェントがシグナリングを有効に行うことでマルチエージェント学習の問題を解消することを期待できる。本研究の今後の課題は、シグナリングによる市場への影響の微視的な分析を行うことでどのようなシグナリングが有効であるかを調べ、シグナリングをより有効に活用する手段を発見することである。

◇ 参 考 文 献 ◇

- [荒井 01] 荒井 幸代: マルチエージェント強化学習—実用化に向けての課題・理論・諸技術との融合—, 人工知能学会誌, Vol. 16, No. 4, pp. 476-481 (2001)
- [Dominguez 93] Dominguez, K. M. and Frankel, J. A.: *Does Foreign Exchange Intervention Work?*, Institute for International Economics, Washington, DC (1993)
- [Goldberg 89] Goldberg, D. E.: *Genetic algorithms in search, optimization, and machine learning*, Addison-Wesley Publishing Company (1989)
- [Grefenstette 88] Grefenstette, J. J.: Credit Assignment in Rule Discovery Systems Based on Genetic Algorithms, *Machine Learning*, Vol. 3, pp. 225-245 (1988)
- [和泉 00] 和泉 潔, 植田 一博: 人工市場アプローチによる為替シナリオの分析, コンピュータソフトウェア, Vol. 17, No. 5, pp. 47-54 (2000)
- [国際 02] 国際金融情報センター: 市場解説 (1995-2002)
- [松井 02] 松井 宏樹, 東条 敏: 介入エージェントを用いた人工市場アプローチによる介入政策の分析, 第 16 回人工知能学会全国大会 (2002)
- [松井 03a] 松井 宏樹, 永田 裕一, 東条 敏: 人工市場アプローチにおける強化学習を用いた介入政策の分析, 第 131 回情報処理学会知能と複雑系研究会 (2003)
- [Matsui 03b] Matsui, H. and Tojo, S.: Artificial Market with Intervention Agent, in *Proceedings of the 1st Indian International Conference on Artificial Intelligence* (2003)
- [日本 02] 日本経済新聞社: 「金融アウトルック」, 日曜版, 日本経済新聞 (1995-2002)
- [Ramaswamy 00] Ramaswamy, R. and Samiei, H.: The Yen-Dollar Rate: Have Interventions Mattered?, *IMF Working Paper* (2000)
- [渡辺 94] 渡辺 努: 市場の予想と経済政策の有効性, 東洋経済新報社 (1994)
- [吉本 00] 吉本 佳生: ニュースと円相場から学ぶ使える経済学入門, 日本評論社 (2000)
- [財務 04] 財務省: 外国為替平衡操作の実施状況 (2001-2004), <http://www.mof.go.jp/1c021.htm>

〔担当委員: 和泉 潔〕

2004 年 7 月 14 日 受理

—— 著 者 紹 介 ——



松井 宏樹

2000年京都産業大学理学部物理学科卒業。2002年北陸先端科学技術大学院大学情報科学研究科博士前期課程修了。現在、同大学同研究科博士後期課程在学中。マルチエージェントシステムと強化学習に興味を持つ。



東条 敏(正会員)

1981年東京大学工学部計数工学科卒業，1983年東京大学大学院工学系研究科修了。同年三菱総合研究所入社。1986-1988年，米国カーネギー・メロン大学機械翻訳センター客員研究員。1995年北陸先端科学技術大学院大学情報科学研究科助教授，2000年同教授。1997-1998年ドイツ・シュトゥットガルト大学客員研究員。博士(工学)。自然言語の形式意味論，オーダーソート論理，マルチエージェントの研究に従事。情報処理学会，人工知能学会，ソフトウェア学会，言語処理学会，認知科学会，Folli 各会員。