

エージェントベースシミュレーションに おける Q 学習を用いたうわさの 流布状況の解析

中 野 統 英

1. 緒 言

現代社会においては経済活動や社会活動が複雑化しており、これらの活動における結果や効果を予測することが難しくなっている。そこである特定の経済・社会活動（例えば株価の予測など）についてモデルを作成し、それを用いたシミュレーションを行うことによって活動の解析や予測を行う手法が多く行われている¹⁾。こういった経済・社会現象のシミュレーションには様々な方法があるが、経済・社会活動をミクロ的な視点からモデルを作成して解析・予測を行うエージェントベースシミュレーションと呼ばれる手法が近年多く用いられるようになってきた²⁻¹³⁾。エージェントベースシミュレーションとは、仮想空間内に配置されたエージェントが決められたルールに従って意思決定や行動を行うことによって、社会現象などを仮想空間上に再現して解析を行う手法である。

著者は人々の間で起こる流行の流布や衰退といった物事に対する採用行動をこのエージェントベースシミュレーションで解析する研究を行っている⁸⁻¹³⁾。これまでにうわさの重要度やあいまいさに着目した研究⁸⁻¹⁰⁾や、経

過時間による採用者数の変化に着目した解析を行った^{11,12)}。実際の人間社会ではそれぞれの人間が判断する時々の環境に適応して最適な行動をとることが多く、これをエージェントの世界で再現しようとする初期設定した行動ルールや（判断時に現れる）偶然性だけでは環境の変化に適応できるエージェントを実現することは難しい。そこで本研究では周りの環境変化に対応して学習を行うことができるエージェントについて研究を行うことにした。今回はQ学習と呼ばれる強化学習の一手法を導入することより、環境変化に対しても柔軟に変化および対応できるエージェントシステムを構築することを目的とする。

2. モデルの構築

2.1 エージェントと仮想空間

ある新しいうわさが伝わっていくときに、そのうわさを採用する革新的な人と容易には採用しない保守的な人があるのは論をまたない。過去にRogersは採用者カテゴリーを5つに分類した（図1）¹⁴⁾。図1ではInnovatorsがうわさを最初に採用し、時間の経過とともにEarly Adopter, Early Majority, Late Majorityの順で採用してLaggardsが最後に採用する、あるいは最後まで採用しないということを表している。その後、

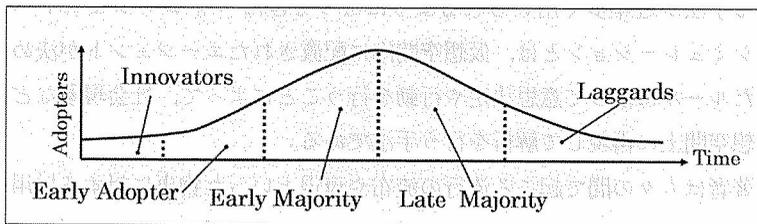


図1 Rogersの提案した採用者カテゴリー

解析を容易にするためにこれらをもとにした2カテゴリーへの分類が提案され^{6,7)} (図2参照), 本研究でもこの2カテゴリーによる分類方法を採用している。ここでは採用者(エージェント)をPioneerとFollowerの2つに分類している。新しい物事の採用に抵抗感が低い革新的なエージェントをPioneerとし, 逆に抵抗感が高い保守的なエージェントをFollowerと定義する。

これらのエージェントが行動を行う仮想空間については, ここでは10×10の格子状空間を考えることにする⁶⁻¹³⁾ (図3参照)。また両端のエー

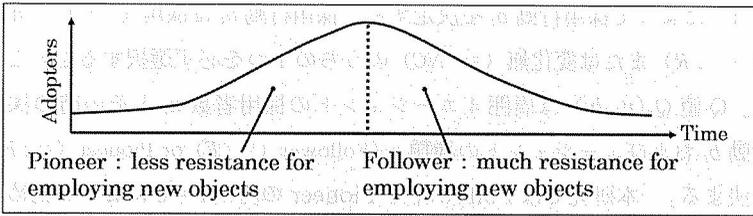


図2 本研究で用いた採用者カテゴリー

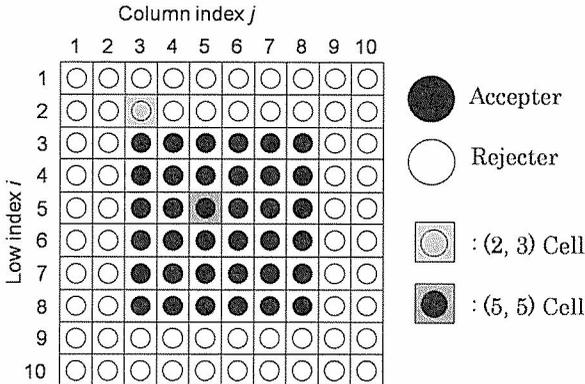


図3 仮想空間とそこに配置されたエージェント

エージェントは反対側のエージェントとつながる構造（トラス構造）であるとする。ここにエージェントを隙間なく配置して予め定めたルールに従って行動させることにより、各エージェントのうわさに対する採用行動（採用（Acceptance）、非採用（Reject）または変化無（No Change））の変化をみるのであるが、本研究では仮想空間中の100エージェントはすべてFollowerかすべてPioneerかのどちらかであるとする。またここで*i*行*j*列目のセルを（*i*, *j*）セルと呼ぶことにする（図3参照）。

図4はエージェントの意思決定プロセスについての説明である。エージェントは時刻*t*のときの周囲4エージェントの採用者数 n_t ($n_t=0, 1, 2, 3$ or 4)によって採用行動 b_t を決定する。採用行動 b_t は採用（=:A）、非採用（=:R）または変化無（=:NC）のうちの1つを必ず選択するものとする。Q値 $Q_e(n_t, b_t)$ は周囲4エージェントの採用者数 n_t とその時の採用行動 b_t およびエージェントの種類*e*（Follower（=:F）or Pioneer（=:P））で決まる。本研究ではFollowerとPioneerの時刻*t*=0における初めの意思決定ルールとQ値を表1および表2のように定義する。

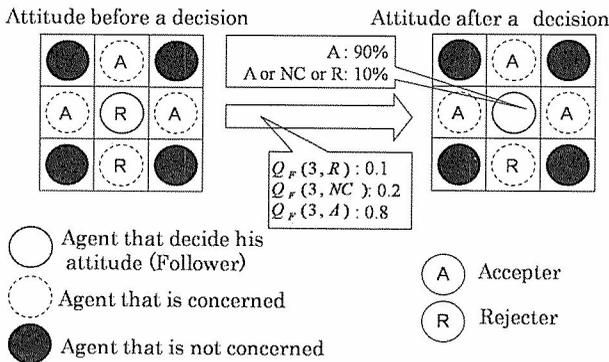


図4 意思決定プロセスの例

表1 Follower の意思設定ルールと Q 値 (初期設定値)

n_t	Initial Rule	Initial values of Q values		
0	R	$Q_F(0, A) = 0$	$Q_F(0, NC) = 0$	$Q_F(0, R) = 1$
1	R	$Q_F(1, A) = 0$	$Q_F(1, NC) = 0$	$Q_F(1, R) = 1$
2	NC	$Q_F(2, A) = 0$	$Q_F(2, NC) = 1$	$Q_F(2, R) = 0$
3	A	$Q_F(3, A) = 1$	$Q_F(3, NC) = 0$	$Q_F(3, R) = 0$
4	A	$Q_F(4, A) = 1$	$Q_F(4, NC) = 0$	$Q_F(4, R) = 0$

表2 Pioneer の意思設定ルールと Q 値 (初期設定値)

N_t	Initial Rule	Initial values of Q values		
0	R	$Q_P(0, A) = 0$	$Q_P(0, NC) = 0$	$Q_P(0, R) = 1$
1	A	$Q_P(1, A) = 1$	$Q_P(1, NC) = 0$	$Q_P(1, R) = 0$
2	NC	$Q_P(2, A) = 0$	$Q_P(2, NC) = 1$	$Q_P(2, R) = 0$
3	R	$Q_P(3, A) = 0$	$Q_P(3, NC) = 0$	$Q_P(3, R) = 1$
4	A	$Q_P(4, A) = 1$	$Q_P(4, NC) = 0$	$Q_P(4, R) = 0$

2.2 意思決定の方法

ここでは図4を用いて意思決定プロセスとQ値の更新手順について説明を行う¹⁵⁾。まずこの図では実線で描かれた中央のエージェントはFollowerでかつ採用者であるとし、このエージェントの採用行動に注目することにする。はじめに破線で示された上下左右の4エージェント中の採用者数をカウントする。ここでは黒丸で示されたエージェントは実線の中央エージェントの意思決定には関係しない。この例の場合、採用者は3つのエージェントとカウントされるので、 $n_t = 3$ の時のQ値をチェックする。この例では、Q値は次のようになっている：

$$Q_F(3, R) = 0.1, Q_F(3, NC) = 0.2 \quad \text{and} \quad Q_F(3, A) = 0.8.$$

この例の場合、 $Q_F(3, A)$ が3つのうちで一番大きな値となっている。本研究では行動決定において ϵ -greedy手法($\epsilon = 0.1$)を用いることにしているので、90%の確率で態度を“採用”に変化させることとなり、残り

の10%の確率で“採用”，“非採用”，“変化無”のうちのどれかの態度をランダムに選択することとなる。

2.3 Q値の更新

Q値 $Q_c(n_t, b_t)$ の更新は時刻ステップ t における意思決定が行われた後に以下のアルゴリズムによって行われる：

$$Q_c(n_t, b_t) \leftarrow Q_c(n_t, b_t) + \alpha \left[r + \gamma \max_{b_{t+1} \in B(n_{t+1})} Q_c(n_{t+1}, b_{t+1}) - Q_c(n_t, b_t) \right], \quad (1)$$

ここで $Q_c(n_{t+1}, b_{t+1})$ は1時刻先に周囲の状態が n_{t+1} になった場合の採用行動 b_{t+1} に対応するQ値である。 r は状態 n_t から n_{t+1} へ遷移において得られる報酬， B_{t+1} は状態 n_{t+1} で実行可能な行動全体の集合を表し， α ($0 < \alpha \leq 1$) は学習率， γ ($0 \leq \gamma \leq 1$) は割引率を表している。Q値の更新はエージェントが一つの行動を実行し、次の状態に移るごとに行われる。

3. シミュレーションの方法

これまでの議論に従い、本研究では以下の方法によってコンピュータ・シミュレーションを行うことにする。なお数値解析についてはJAVA言語を用いてプログラム作成を行った。

- [1] 表3に示した7パターンのシミュレーションのうちの1つを選択して各種のパラメータを設定する。
- [2] 100個のエージェント（すべてFollowerまたはすべてPioneer）を 10×10 の（トーラス状）格子空間上に配置し、初期採用者数および配置を与える。なお本研究では格子空間の中央部分に正方形に初期採用者を配置するものとする（図3参照）。
- [3] FollowerとPioneerのQ値 $Q_c(n_t, b_t)$ の初期値を表1と表2に示したように定義する。

表3 シミュレーションパターンと解析結果

No.	Agent	Initial Acceptor	ϵ -greedy	Q-learning	A_{ave}
1	Follower	36 (6 × 6)	○	○	868.8
2	Follower	36 (6 × 6)	○	×	1.5
3	Follower	36 (6 × 6)	×	×	(0)
4	Pioneer	36 (6 × 6)	○	○	0.0
5	Follower	16 (4 × 4)	○	○	909.7
6	Follower	25 (5 × 5)	○	○	895.8
7	Follower	49 (7 × 7)	○	○	744.3

- [4] 各エージェントは $Q_e(n_i, b_i)$ に基づいて1時刻ステップにつき1回態度を決定し、それから(1)式を計算して $Q_e(n_i, b_i)$ を更新する(図4参照)。
- [5] すべてのエージェントが採用状態になるまで解析を行い、この時点で報酬 r を付与する。もし全エージェントが採用状態にならなければ1000ステップ ($t=1000$) で解析を強制終了させる。
- [6] 上記のプロセス [4] と [5] を1000回繰り返し、この中での終了回数(ここでの終了とはすべてのエージェントが1000ステップ以内に採用状態になることとする)をカウントする。これを1試行とする。
- [7] 1つの条件(表3参照)について10試行を行い、このときの収束回数の平均値(=: A_{ave})を計算する。

4. 解析および考察

4.1. 今回提案したモデルの効果

シミュレーションの結果 (A_{ave}) を表3に示す。本研究では Simulation No. 3以外については ϵ -greedy 手法 ($\epsilon=0.1$) を用いた。Simulation No. 3については greedy 手法 (ϵ -greedy 手法において $\epsilon=0$) を用いている。あと学習率は $\alpha=0.1$ 、割引率は $\gamma=0.9$ および報酬は $r=10$ と設定した。

表3の解析結果 A_{ave} において Simulation No. 1, 2 および 3 を比較すると、Q 学習を導入しない場合 (No. 2 および 3) はほとんどシミュレーションが終了していないが、導入した場合 (No. 1) は約 87% というかなり高い確率で終了していることが分かる。この結果によりエージェントベースシミュレーションに Q 学習を導入すると効果が現れる場合があることが実証された。なお No. 2 において A_{ave} が 1.5 回となっているが、これは確率的行動選択手法である ϵ -greedy 手法を用いているためにまれに収束する場合があるためだと思われる。

また表3の No. 1 と 4 を比較すると、Follower の場合はかなりの割合で終了しているのに対して Pioneer の場合はシミュレーションを終了できない (全エージェントが情報を採用している状態にならない) ことが分かる。これは Pioneer の初期ルールの設定による所が大きい (とくに $n_i=3$ で非採用とするルールになっている影響が大きい) と思われる。よって採用行動の状況はエージェントの初期行動選択ルールの設定に大きく影響を受けることが分かる。

さらに表3の No. 1 と 5 から 7 までを比較するとどの条件でも 7 割以上の割合で終了しており、Follower によるシミュレーションにおいては初期採用者の数にかかわらず Q 学習手法がうまく機能していることが分かる。ただ初期採用者が少ないほうが A_{ave} が多くなる傾向があるのは、 ϵ -greedy 手法を用いている関係で初期の非採用者が多い場合に非採用者が採用者に突然変化する影響が大きく出ているものと思われる。

4. 2. 採用者数の時間的变化

ここでは採用者数の時間経過に対する推移をみることにする。図5は表3における Simulation No. 1 の時の採用者数の時間的推移であり、図6および7は Simulation No. 7 および 2 のときの推移である。これらの図ではいずれも 100 回のシミュレーション結果が描かれており、時刻 $t=t_{end}$

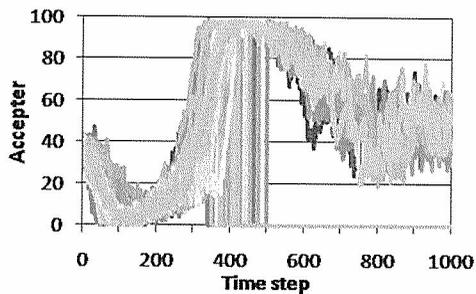


図5 Simulation No. 1における採用者数の推移

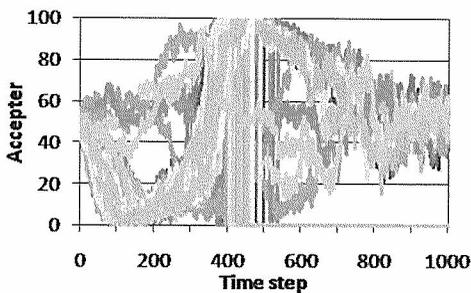


図6 Simulation No. 7における採用者数の推移

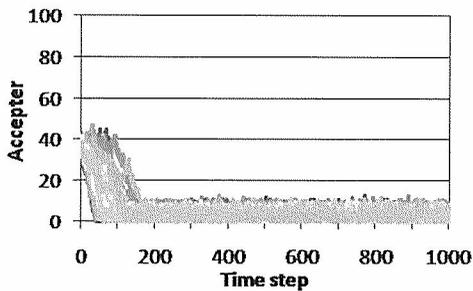


図7 Simulation No. 2における採用者数の推移

で採用者数が100になった場合（シミュレーションが終了した場合）は時刻 t_{end} 以降 ($t > t_{end}$) の採用者数を強制的に0としている。

図5の Simulation No. 1 においては、10回のシミュレーションで終了しなかった。90回のシミュレーションでは時刻 $345 \leq t \leq 500$ の間で終了した。この時の変化を解析すると、 $0 \leq t \leq 200$ でいったん採用者数が減少してから徐々に増加する傾向がある。図6の Simulation No. 7 では、12回のシミュレーションにおいて終了しなかった。88回のシミュレーションでは時刻 $374 \leq t \leq 537$ の間で終了しているが、上述の傾向に従わないランダムな挙動を示すシミュレーションがいくつか現れている。図7の Simulation No. 2 においては $t \leq 1000$ で終了したシミュレーションはなく、ほとんどがランダムな挙動を示している。この場合の初期採用者数は36であるが、 $t \geq 200$ で採用者数が20以下になっている。

4. 3. Q 値の時間的变化

最後に Q 値の挙動について解析を行った。本論文では Simulation No. 1 における (5, 5) セルに配置されたエージェントの Q 値に着目してその推移を解析した。 $Q_F(0, A)$ 、 $Q_F(0, NC)$ および $Q_F(0, R)$ の推移をそれぞれ図8、9および10に示した。同様に $n_i=1, 2, 3$ および4における Q_F

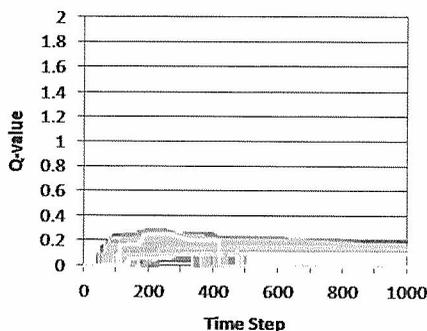


図8 Simulation No. 1 における $Q_F(0, A)$ の推移

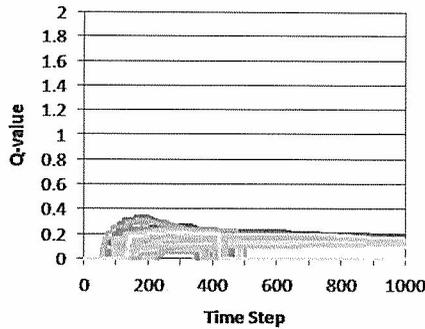


図9 Simulation No. 1における $Q_F(0, NC)$ の推移

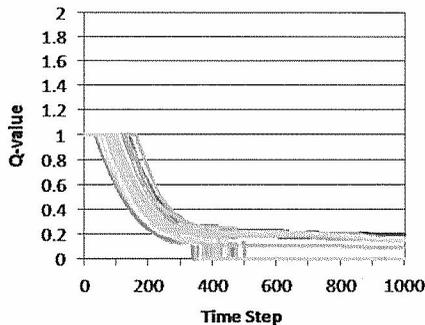


図10 Simulation No. 1における $Q_F(0, R)$ の推移

(n_t, A) , $Q_F(n_t, NC)$ および $Q_F(n_t, R)$ を図11から22にまとめて示した。これらの図でも1つの図につきそれぞれ100回のシミュレーションにおけるQ値の挙動が描かれており、時刻 $t = t_{end}$ で採用者数が100回になった場合（シミュレーションが終了した場合）は時刻 t_{end} 以降 ($t > t_{end}$) のQ値も強制的に0にリセットされる設定としている。

$n_t = 0, 1$ および 2 のときのQ値はよく似た挙動を示している。 $Q_F(0, R)$, $Q_F(1, R)$ および $Q_F(2, NC)$ という初期値が1のQ値は減少傾向を示している。また他の初期値が0のQ値では増加傾向を示しているものの値

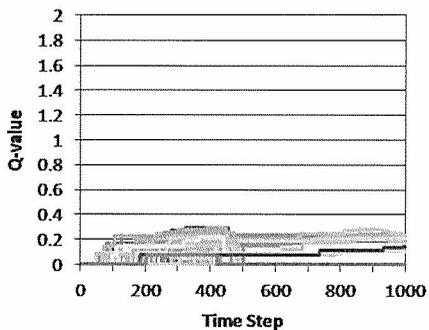


図 11 Simulation No. 1 における $Q_F(1, A)$ の推移

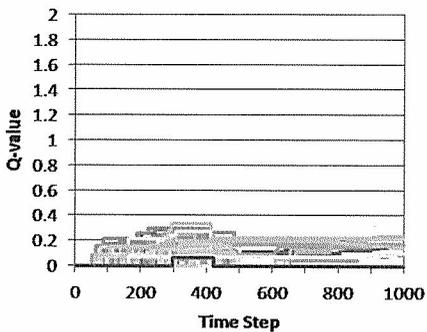


図 12 Simulation No. 1 における $Q_F(1, NC)$ の推移

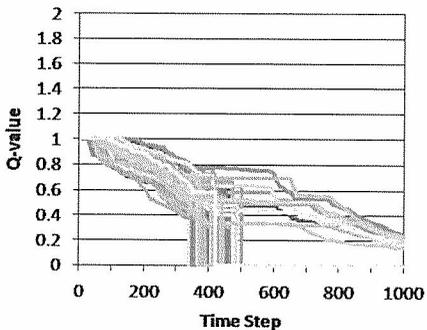


図 13 Simulation No. 1 における $Q_F(1, R)$ の推移

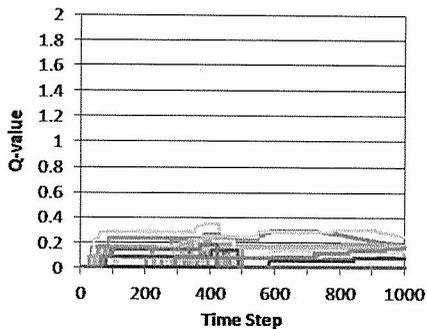


図 14 Simulation No. 1 における $Q_F(2, A)$ の推移

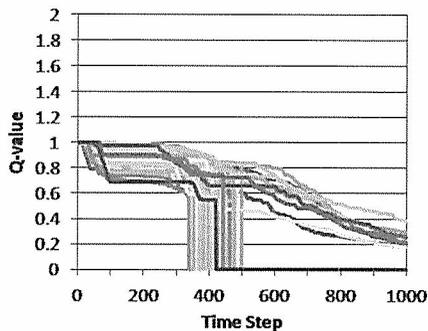


図 15 Simulation No. 1 における $Q_F(2, NC)$ の推移

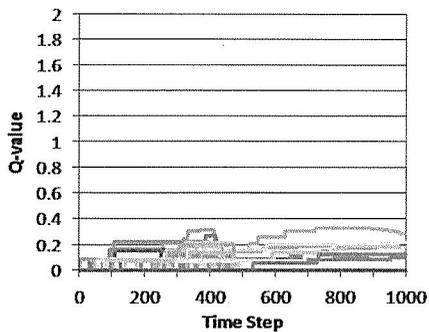


図 16 Simulation No. 1 における $Q_F(2, R)$ の推移

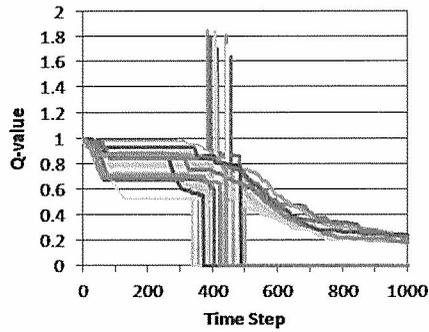


図17 Simulation No. 1における $Q_F(3, A)$ の推移

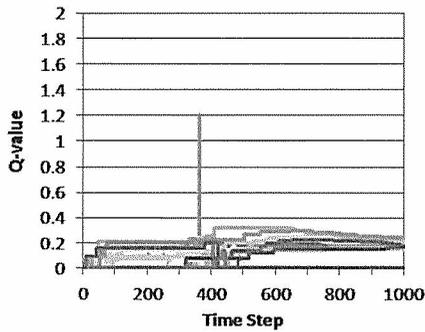


図18 Simulation No. 1における $Q_F(3, NC)$ の推移

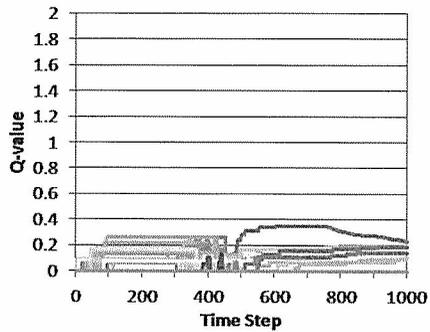


図19 Simulation No. 1における $Q_F(3, R)$ の推移

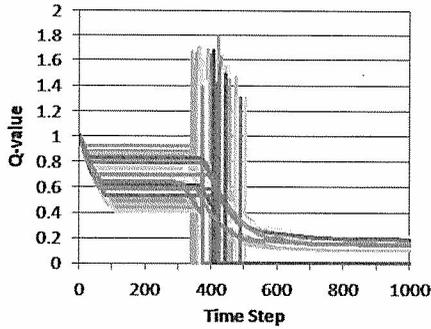


図 20 Simulation No. 1 における $Q_F(4, A)$ の推移

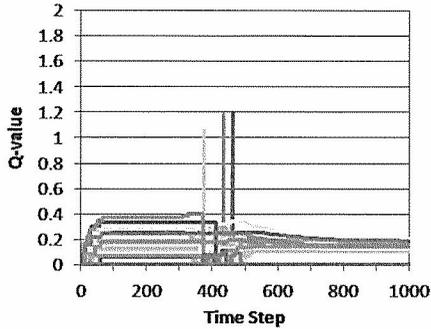


図 21 Simulation No. 1 における $Q_F(4, NC)$ の推移

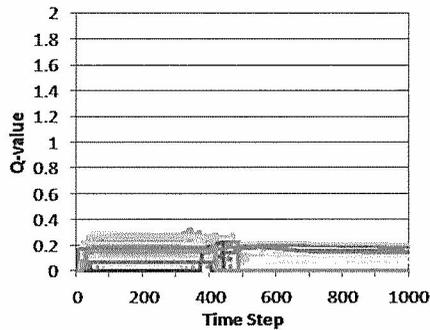


図 22 Simulation No. 1 における $Q_F(4, R)$ の推移

はすべて 0.4 以下のままになっている。

$n_t = 3$ および 4 のときの Q 値も同じような挙動を示している。 $Q_F(3, A)$ および $Q_F(4, A)$ といった初期値が 1 の Q 値は減少傾向を示している。また他の初期値が 0 の Q 値では増加傾向を示しているものこのでも値はすべて 0.4 以下のままになっている。図 17, 18, 20 および 21 でパルス状の挙動が出ている Q 値があるが、これらはシミュレーションが終了して報酬を得るため急激な上昇を示しているからである。

5. 結 言

本研究ではエージェントベースシミュレーションでのエージェントに Q 学習アルゴリズムを導入することにより、学習効果をもつエージェントの有用性を立証した。本研究で得られたポイントは、意思決定ルールに時々刻々変化させることのできる Q 値を導入しこれらを効果的に変化させることによってエージェントの性格を柔軟に変化させることができることである。

ただ Q 学習手法における各種パラメータの設定、特に学習率と割引率についてはこれからさらなる検討を要する。さらにエージェントにおける行動ルールの設定や採用行動の収束性についての解析など多くの課題があるが、これからはエージェントベースシミュレーションへの強化学習手法の導入を積極的に推進していきたい。

参考文献

- 1) 石村貞夫, 石村園子: 金融・証券のためのブラック・ショールズ微分方程式, 東京図書 (1999).
- 2) 喜多一: エージェントベースアプローチのための情報系要素技術, 計測自動制御学会誌, Vol. 43, No. 12, pp. 938-943 (2004).

- 3) 山影進, 服部正太編: コンピュータのなかの人工社会 マルチエージェントシミュレーションモデルと複雑系, 共立出版 (2002).
- 4) J. M. Epstein and R. Axtell (服部正太, 木村香代子訳): 人工社会 —— 複雑系とマルチエージェント・シミュレーション ——, 共立出版 (1999).
- 5) 出口弘, 倉田正, 田沼英樹: エージェントベースの社会シミュレーション言語 SOARS の設計思想とその展開, 計測自動制御学会 システム・情報部門 第 35 回システム工学会研究会「エージェントベース社会システム科学の創出」～新しい学問領域の確立を目指して～資料, pp. 153-158 (2005).
- 6) 藤井慎太郎: エージェントベースシミュレーションを用いた社会的拡散現象の解析, 北陸先端科学技術大学院大学知識科学研究科修士論文 (2003).
- 7) 村田忠彦, 八木貴志: エージェントベースシミュレーションにおける初期採用者の分布と採用行動の伝播の関係, 第 48 回システム制御情報学会研究発表講演会公演論文集, pp. 177-178 (2004).
- 8) 中野統英: 社会的拡散現象における重要度を考慮した影響力の解析～エージェントベースシミュレーションを用いて～, 日本繊維機械学会誌 Vol. 61, No. 9, pp. 615-620 (2008).
- 9) N. Nakano: Analysis of Influences Considering Vagueness in Agent-Based Simulations, Proc. 39th ISIC International Symposium on Stochastic Systems Theory and Its Applications (SSS '07), pp. 115-120 (2008).
- 10) N. Nakano: Influences of Rumors Including Vagueness in Agent-Based Simulations, Trans. International Journal of Innovative Computing, Information and Control (IJICIC), Vol. 5, No. 1, pp. 129-138 (2009).
- 11) 中野統英: エージェントベースシミュレーションにおける前期採用者の採用ルールの違いによる影響の解析, 第 61 回日本繊維機械学会年次大会公演論文集, pp. 120-121 (2008).
- 12) 中野統英: エージェントベースシミュレーションにおける前期採用者の採用ルールの違いによる採用者数の変化の解析, 追手門経営論集, Vol. 14, No. 2, pp. 93-105 (2008).
- 13) 中野統英: マルチエージェントシミュレーションにおけるうわさの流布状況の解析 —— エージェントシミュレーションへの Q 学習の導入 ——, 第 62 回日本繊維機械学会年次大会公演論文集, pp. 74-75 (2009).
- 14) E. M. Rogers (青池慎一, 宇野善康訳): イノベーション普及学, 産能大

学出版部 (1990).

- 15) 高玉圭樹：マルチエージェント学習 —— 相互作用の謎に迫る ——, コロナ社 (2003).