

学習エージェントを導入した人工市場における均衡価格の検証

兼平 大輔[†] 川村 秀憲[†] 車谷 浩一[‡] 和泉 潔^{‡‡} 大内 東[†]

[†]北海道大学大学院工学研究科 〒060-8628 北海道札幌市北 13 条西 8 丁目

^{‡, ‡‡}産業技術総合研究所 サイバーアシスト研究センター 〒135-0064 東京都江東区青海 2-41-6

E-mail: [†] {kanehira, kawamura, ohuchi}@complex,eng.hokudai.ac.jp,

[‡] kurumatani@w-econ.org, ki, ^{‡‡} kiyoshi@ni,aist.go.jp

あらまし 近年, エージェントベース人工市場モデルの研究が多数行われている, しかし, これらの多くの研究で構築する市場のモデルは板寄せ方式のモデルか, ザラバ方式のモデルのどちらか一方のみであった.

そこで, 本稿では, 市場モデルとして板寄せ方式の市場とザラバ方式の市場の両方を構築し, その中にトレーダーとして学習エージェントを投入しその市場の挙動の検証を行う.

実験において需要曲線と供給曲線から得られる理論的な均衡価格と実際に得られる約定価格とを比較した結果, 学習トレーダーによって構築される市場において, 板寄せ方式を採用している方がザラバ方式よりも実際の約定価格が均衡価格に近づくということが判明した.

キーワード 需要曲線, 供給曲線, マルチエージェントシステム, X-Economy System

Verification of the Equilibrium Price in the Artificial Market which introduced the Learning Agent

Daisuke KANEHIRA[†] Hidenori KAWAMURA[†] Koichi KURUMATANI[‡]

Kiyoshi IZUMI^{‡‡} and Azuma OHUCHI[†]

[†] Graduate School of Engineering, Hokkaido University

Kita 13 Nishi 8, Kita-ku, Sapporo, 060-8628 Japan

[‡] Cyber Assist Research Center (CARC),

National Institute of Advanced Industrial Science and Technology (AIST)

Aomi 2-41-6, Koto-ku, Tokyo, 135-0064 Japan

E-mail: [†] {kanehira, kawamura, ohuchi}@complex,eng.hokudai.ac.jp,

[‡] kurumatani@w-econ.org, ki, ^{‡‡} kiyoshi@ni,aist.go.jp

Abstract Recent years. Much researchers study an agent base artificial market model. However whose model of the market built by research of these many was either the model of itayose system, or a model of a zaraba system. In our experiment which builds both the market of itayose system, and the market of a zaraba system as a market model. Supplies a study agent as a trader in it in this paper, and verifies the action of the market As a result of comparing the theoretical equilibrium price obtained from a demand curve and a supply curve with the agreement price actually obtained, in the market built by the study trader, it became clear that an agreement price with more nearly actual than a zaraba system having adopted the itayose system approaches an equilibrium price.

Keyword demand curve, supply curve, multi-agent system, X-Economy System

1. はじめに

近年, エージェントベース人工市場モデルの研

究が多数行われている, しかし, これらの多くの研究で構築する市場のモデルは注文を一度市場

に集約し、集まった注文を付合わせ一斉に売買の締結を行う板寄せ方式のモデルか、市場にきた注文を一つずつ順番に付合わせて売買の締結を行うザラバ方式のモデルのどちらか一方のみであった。

そこで、本稿では、市場モデルとして板寄せ方式の市場とザラバ方式の市場の両方を構築し、その中にトレーダーとして学習エージェントを投入しその市場の挙動を比較する。

2. 人工市場モデル

本稿で用いる人工市場モデルは、独自の戦略を持ちその戦略に基づく需要・供給を注文という形で市場に反映させるトレーダーエージェントと、そのエージェントからの注文を集約し売買を締結させる市場から構成されるマルチエージェントシステムになっている。

この人工市場では、1ステップ単位で時間が進み、1日のステップ数は固定されている。

また、本稿では X-Economy System[1][2]を用いて、人工市場モデルをコンピューター上に実装した。

2.1. 市場の売買締結方法

通常市場では、価格優先の原則と時間優先の原則に従い、売り注文間の競争と買い注文間の競争を行い、最も優先する売り注文と最も優先する買い注文が値段的に合致した時に、その値を約定値段として売買を成立させる。価格優先の原則では、売りについては値段の低い注文が、買いについては値段の高い注文が優先する。

また、時間優先の原則では、同値の注文については、先に出された注文が後に出された注文に優先する。

本稿ではこの優先原則を用いた板寄せ方式とザラバ方式の2種類の売買締結方式を実装した市場を用いる。

2.1.1. 板寄せ方式

板寄せ方式とは、価格優先の原理に従い売注文と買注文のバランスに応じて売買を成立させていく売買締結方法である。具体的には、値段決定前の注文をすべて注文控え(板)に記載したうえで、価格の優先順位の高い注文から順次対応させ、数量的に合致する値段を決定し、その価格を単一の約定値段とする。

この方式は、取引所において、寄付きと呼ばれる売買立会の始めの約定値段である始値と引けと呼ばれる売買立会終了時の約定値段である終値を決定する場合に用いられている。

2.1.2. ザラバ方式

ザラバ方式とは、注文を時間優先、価格優先の原則に

従いながら順次執行していき、すでに注文されているもっとも優先される注文と組み合わせる売買を締結させる方法である。注文価格以下(以上)の売り(買い)の注文がまだなかった場合の買い(売り)注文は市場に残り約定を待つことになる。

また、この方式は取引所において、ザラバと呼ばれる寄付きと引け以外の売買立会中に継続して用いられる方式である。

2.2. トレーダー

本モデルのトレーダーはあらかじめ売り手(seller)と買い手(buyer)に分かれており、売り手が買い手に、買い手が売り手にその立場を変えることはない。

トレーダーの目的は、売り手ならば、1日にある商品を1単位をより高く売ることであり、逆に買い手ならばその商品1単位をより安く買うことである。

また、トレーダーは各々取引のための戦略を持っており、その戦略に基づき各ステップで注文価格を決定し、市場に1単位の売買の注文をする。

市場に出した注文が約定する、あるいは、1日の全てのステップが終わってしまうとトレーダーはその日の行動を終了する、全トレーダーの行動が終了すると、市場は次の日に移行し、再び全てのトレーダーが取引を成立させようと注文を行いはじめる。

トレーダーが各ステップに注文する価格は shout と表される、トレーダーはこの shout を各々のトレーダーに与えられた基準価格(base)とトレーダーが手に入れた利益率(profit)を用いて以下のように算出する、

そのトレーダーが売り手(Seller)の場合

$$\text{shout} = \text{base} + (\text{profit} \cdot \text{base})$$

(1)

また、買い手(Buyer)の場合

$$\text{shout} = \text{base} - (\text{profit} \cdot \text{base})$$

また本稿で用いられるトレーダーはこの profit を算出する戦略の違いにより以下の5種類に分類できる。

- Zero Intelligence under Unconstraint (ZIU) トレーダー
- Zero Intelligence under Constraint (ZIC) トレーダー
- Simple Intelligence under Unconstraint (SIU) トレーダー
- Simple Intelligence under Constraint (SIC) トレーダー
- Q-Learning Constraint (QLC) トレーダー

ZIU, ZIC は注文価格を算出するための profit をラン

ダムに決定する (Zero Intelligence) という戦略をとる,

SIU,SIC が持つ戦略は1日の取引の中で,もし注文が約定しなかった場合は1つ前のステップの注文よりも小さな profit をランダムに決定するという単純な戦略 (Simple Intelligence) であり,これは売り手 (買い手) にとって,取引が成立しなかった場合は次のステップでは注文価格を下げる (上げる) 事により,取引を成立させやすくするという戦略である,

また,C (Constraint),U (Unconstraint) の違いであるが,これはトレーダー個別に与えられるに予算の制約 (Constraint) があるかないかの区別である,この予算とは,例えば売り手にとっては入れ価格,買い手にとっては保有する所持金などがあてはまり,その価格より安く (高く) 売る (買う) ことができないという最低 (最高) 価格を表すものでもある,

本モデルにおいてトレーダーは市場という度か注文を出した後に取引の成立にたどり着くような報酬の遅れが存在する環境の中にいるので,試行錯誤的に行動し時々環境から与えられる報酬をもとに行動を修正していく学習手法である強化学習の一つの Q-Learning[3]を用いて,学習エージェントは学習を行う,

この Q-Learning において状態 s_t において行動 a_t ,

$$\hat{Q}(s_t, a_t) \leftarrow \hat{Q}(s_t, a_t) + \alpha \{r_t + \gamma \max_{a \in A} \hat{Q}(s_{t+1}, a) - \hat{Q}(s_t, a_t)\}$$

をとって状態 s_{t+1} に遷移し,報酬 r_{t+1} が得られたとき Q 値の更新式は以下のとおりになる,

この式において α は学習率, γ は割引率であり,現在の状態においてある行動をとったときの Q 値をその次状態において Q 値が最も高い値に近づけるに変更することを意味している,

また,本稿では学習エージェントは行動 a_t を,可能な行動の中から確率 ε でランダムに選択し,確率 $(1-\varepsilon)$ で Q 値のもっとも大きくなる行動 a_t を選ぶという, ε グリーディ方策を採用した,

本稿では Q-learning トレーダーのパラメータを表 2 のように設定した,

α	0,5
γ	0,5
s_t	注文回数
a_t	profit の選択(50~0 の 51 通り)
r_t	約定した場合その注文に用いた profit 約定しなかった場合は 0
ε	0,3

表 1:Q-learning のパラメータ,

本稿では,Constraint なトレーダーである ZIC,SIC, QLC の予算には基準価格 (base) を用いている,

基準価格がトレーダーの注文価格の限界価格となり,売り手にとっては限界価格 (=基準価格) が注文価格の下限値となり,買い手にとっては上限値となる,ゆえに (1) より Constraint なトレーダーの profit の最小値は 0 になる,

Trader	Profit の決定方法,
ZIU	+50~-50 の間のランダム値,
ZIC	+50~0 の間のランダム値,
SIU	前注文~-50 の間のランダム値,
SIC	前注文~0 の間のランダム値,
QLC	強化学習 (Q-Learning) を用いる,

表 2:各 Trader の profit の決定方法,

設定	S min	S max	D min	D max
1	75	325	75	325
2	200	200	75	325
3	50	50	200	200
4	200	200	320	320

3. 需要と供給

予算が適応されたトレーダーにより構成される市場は,潜在的な需要と供給の曲線を導くことができる,今,売り手 (A,B,C,D,E) と買い手 (F,G,H,I,J) の各々 5 名ずつ市場に参加しており,おり,A=150,B=200,C=250,D=300,E=350 という予算の制約 (下限) が与えられたとすると図 1 の供給曲線を描くことができる,この供給曲線は 150 未満の価格で売ってくれる売り手はいないこと,150 以上 200 未満の価格では A にだけ売る意思があること,200 以上 250 未満では A,B に売る意思があること,さらに,350 以上の価格では A~E の売り手全員に売る意思があることを示している,

同様に,買い手に F=150,G=200,H=250,H=300, I=350 という予算の制約 (上限を) を与えると図 1 の需要曲線を描くことができる,この需要曲線は,価格が 150 以下であれば買い手である 5 人全員に買う意思があり,価格が上がるにつれて,買い手が減っていくことを表している,

また,このときの均衡価格 P_0 は 250 となり,均衡需給量 Q_0 は 3 となる,

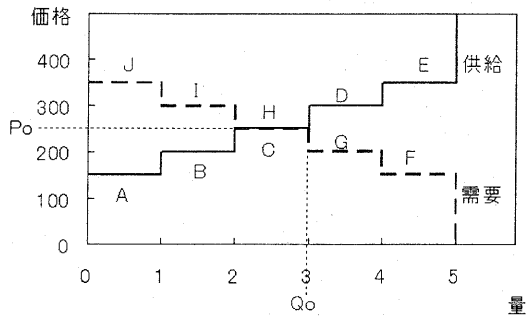


図 1 : 5人ずつの売り手と買い手がある市場の需要供給曲線の例, 均衡価格 P_0 , 均衡需給量 Q_0

4. 実験設定

本稿では,表3のように変化させた4つの設定の基準価格(限界価格)を5種類のトレーダー(ZIU,ZIC,SIU,SIC,QLC)に与え,さらに2種類の売買締結方式(板寄せ,ザラバ)で実験を行った,

表 3 : S_{num}, B_{num} は売り手と買い手の人数
 S_{min}, S_{max} は供給価格の最小値と最大値,
 D_{min}, D_{max} は需要価格の最小値と最大値,

表3に基づいて,需要供給曲線のグラフをかくと図3になる,このグラフからわかるように設定1,2,3,4の全てにおいて需要曲線と供給曲線が交わる点であ

$$s_t, a_t) \leftarrow \hat{Q}(s_t, a_t) + \alpha \{r_t + \gamma \max_{a \in A} \hat{Q}(s_{t+1}, a) - \hat{Q}(s_t, a_t)\}$$

る均衡価格 P_0 は 200 であり,また取引量は6単位になる,

また実験は,1日のステップ数は10ステップとし,1試行を1000日として実験を行った,

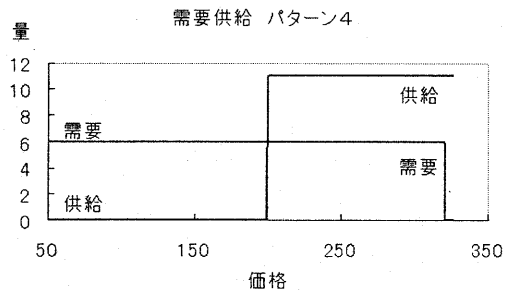
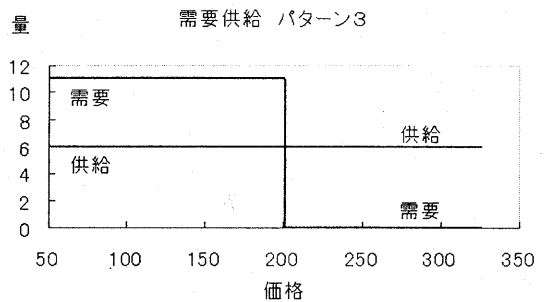
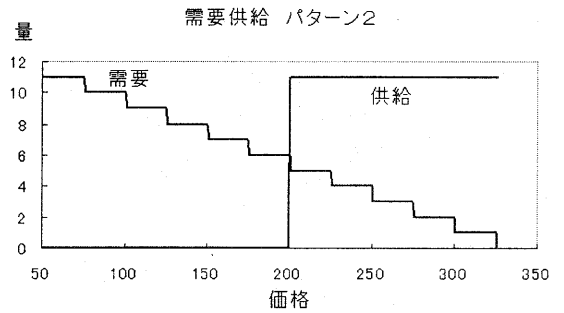
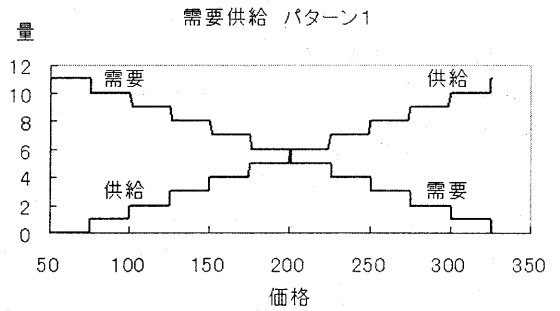


図 2 : 実験設定 1, 2, 3, 4 の需要供給曲線のグラフ

5. 実験結果と考察

表 4 から表 7 は実験で得られた平均価格とその分散を表している,

板寄せ方式とザラバ方式の約定価格の分散の平均を示している表 5 と表 7 から, ザラバ方式のほうが板寄せ方式よりも大幅に分散が大きいことが見て取れる, このことは, 板寄せ方式の市場のほうが価格の安定性が高いことを示している,

また, 同じ表 5, 表 7 から ZIU トレーダーの分散も売買締結方式, 設定にかかわらず他のトレーダーよりも大きくなっている事がわかる, SIU や SIC トレーダーは, profit を減少させていく傾向がある戦略をとっており 1 日の中でステップ数に比例して注文価格が基準価格に収束していくので, ZIU よりも分散が小さくなっていると思われる, また, ZIC トレーダーの profit は (0 ~ 5 0) の間に限定されているので, ZIU ほどは分散が大きくならずにすんでいる,

	設定 1	設定 2	設定 3	設定 4
ZIU	167,3	144,7	190,1	160,9
ZIC	198,9	208,2	145,2	217,1
SIU	190,3	198,3	145,0	220,0
SIC	197,0	204,5	145,3	216,2
QLC	200,5	202,0	199,9	208,3

表 4 : 板寄せ方式における各トレーダーの約定価格の平均

	設定 1	設定 2	設定 3	設定 4
ZIU	341,3	697,7	181,9	1398,8
ZIC	171,0	42,1	43,6	125,7
SIU	65,4	201,3	44	123,7
SIC	57,7	46,1	44,7	152,7
QLC	96,3	13,7	58,3	65,6

表 5 : 板寄せ方式における各トレーダーの約定価格の分散の平均

	設定 1	設定 2	設定 3	設定 4
ZIU	198,2	182,2	196,1	195,0
ZIC	188,8	216,6	145,1	228,2
SIU	187,4	204,7	146,1	227,1
SIC	192,0	213,3	145,1	225,5
QLC	191,7	212,7	144,9	226,7

表 6 : ザラバ方式における各トレーダーの約定価格の平均

	設定 1	設定 2	設定 3	設定 4
ZIU	5375,1	3270,1	6975,5	6047,8
ZIC	1980,1	276,8	2101,5	524,3
SIU	1541,2	456,9	2053,0	562,5
SIC	1413,4	276,0	2106,8	563,6
QLC	1287,7	172,6	1941,4	417,8

表 7 : ザラバ方式における各トレーダーの約定価格の分散の平均

次に表 4, 表 6 の約定価格の平均値を見つめる前に約定価格の期待値 $E(p)$ というものを導入する, Cliff[4]

によると設定 1, 2, 3, 4 の需要供給曲線を持つ ZIC トレーダーで構成される市場の約定価格はそれぞれ図 3, 1, 3, 2, 3, 3, 3, 4 に示される確率密度関数で表される,

この確率密度関数を用いて約定価格の期待値を求める, 設定 1, 2, 3, 4 の約定価格の期待値をそれぞれ $E_1(p), E_2(p), E_3(p), E_4(p)$ とし需要曲線の最大値

を D_{\max} , 供給曲線の最小値を S_{\min} すると,

$$E_1(p) = P_0$$

$$E_2(p) = P_0 + \frac{1}{3}(D_{\max} - P_0)$$

$$E_3(p) = \frac{1}{2}(P_0 + S_{\min})$$

$$E_4(p) = \frac{1}{2}(P_0 + D_{\max})$$

という式で表される,

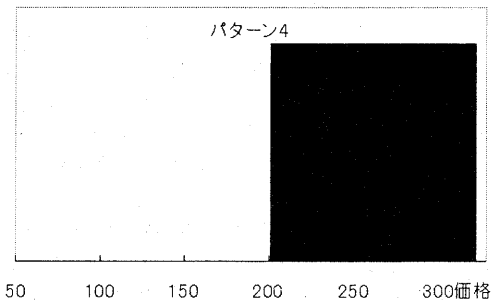
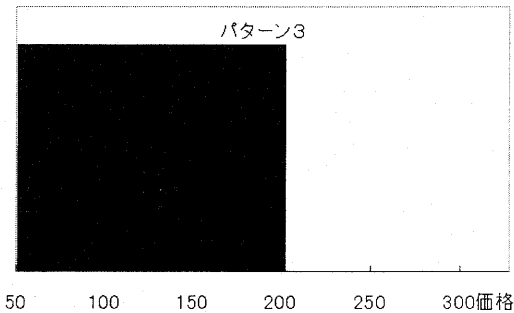
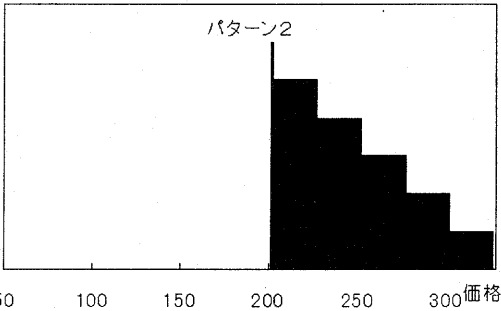
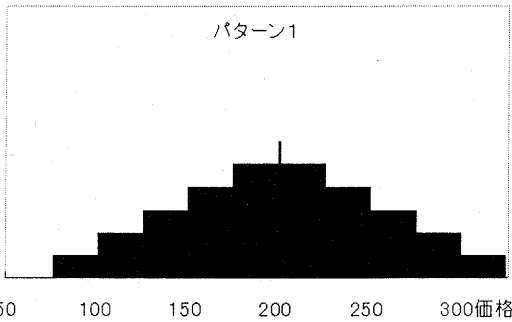
これに具体的な数値を代入すると, 設定 1 = 200, 設定 2 = 233,3, 設定 3 = 125, 設定 4 = 260 となる,

Cliff のモデルではトレーダーの注文価格そのものを, 一様乱数によって決定する手法が用いられているのに対して, 本稿では式 1 のように一様乱数を利益率とみなして注文価格を算出する手法を用いているが, それでも, 表 6 のザラバ市場での約定価格の平均値には Cliff のモデルと同様の傾向が見られ, Q-Learning を用いた QLT トレーダーの平均価格も需要供給曲線の均衡価格から乖離している,

しかし, 表 4 の板寄せ方式における各トレーダーの約定価格の平均を見つめると, 他のトレーダーの約定価格の平均値が均衡価格から約定価格の期待値の方へ乖離している中で, QLC トレーダーの約定価格だけが均衡価格付近に収束している,

学習トレーダーによって構築される市場において, 板寄せ方式の売買締結方式を採用している市場の方が均衡価格が約定価格になりやすく, ザラバ方式の売買

締結方式を採用している市場のほうが均衡価格から乖離した値を約定価格としやすいことが判明した,



6. おわりに

本稿では,X-Economy System を用いて,ザラバ方式と板寄せ方式の市場を構築し,そこに ZIU,ZIC,SIU,SIC,と学習エージェントである QLC トレーダーの5つの戦

略の異なるトレーダーエージェントを投入し,実験を行った結果,学習トレーダーによって構築される市場において,板寄せ方式の売買締結方式を採用している市場の方が均衡価格が約定価格になりやすく,ザラバ方式の売買締結方式を採用している市場のほうが均衡価格から乖離した値を約定価格としやすいことが判明した.

文 献

- [1] 川村 秀憲,車谷 浩一,大内 東,"X-Economy System の紹介",エージェント合同シンポジウム(JAWS 2002),pp, 423-424 (2002),
- [2] X-Economy project official home page, <http://www.x-econ.org>
- [3] Richard S. Sutton and Andrew G. Barto, 強化学習, 三上 貞芳, 皆川 雅章(訳), 森北出版株式会社, (2000), 東京,
- [4] D. Cliff, Minimal-Intelligence Agents for Bargaining Behaviors in Market-Based Environment, technical report HPL-97-91, Technical report, Hewlett Packard Labs, 1997,