

人工経営者の協調戦略学習

佐藤亮

rsato@shako.sk.tsukuba.ac.jp
筑波大学 社会工学系
郵便番号305 つくば市天王台1-1-1

経営をシミュレーションすることを通じて概念スキルを獲得することを目的としたビジネスゲームは、通常は人間プレーヤーによって行なわれる。うまい経営をするための意思決定のルールを獲得するのは、経営環境の不確かさや多くの経営指標で表わされる多目標間の整合のとりかたが複雑であるために、人間であっても困難である。たとえば、単なる期間利益最大化目標からは、製品在庫の拡大・借入金の増大という好ましからざる行動が導かれる。

本研究では、ゲーミング・シミュレーションによって、4プレーヤーが参加する簡単な販売ゲームに対する遺伝的アルゴリズムにもとづくクラシファイアを利用したいく通りかの人工経営者を構成することを通して、多目標の組合せ方とその最大化行動との関係を検討する。また、協調戦略の学習を観察する。

Artificial Managements Learn Cooperative Strategies

Sato Ryo

Institute of Socio-Economic Planning, University of Tsukuba,
Tennnodai 1-1-1, Tsukuba, Ibaraki 305 Japan

Business games simulate business situations under which human players can learn some conceptual skills for business decision. In many cases it is not easy to find good strategic rules for decision in a game, because a player should manage the environmental uncertainty like other players' decisions, and should integrate some goals like profit and R&D investment. A simple minded profit maximizer easily goes to have unnecessary stock of commodities and huge loan.

In this study we apply a classifier systems, that is based on Genetic Algorithm, to get management rules for a simple sales game with four players. We will see what combination of promising goals can lead a set of rules. Also three artificial managements' learning of a kind of cooperative rules will be examined and discussed.

1 はじめに

簡単な販売ゲームについて、遺伝的アルゴリズムにもとづくクラシファイアを利用したいく通りかの人工経営者を構成し、機械学習を通じてルール獲得を行ない、そのルールによるゲーミングシミュレーションを行なう。経営行動ルールの協同現象を観察するためのゲーミング・シミュレーションの構成も行ない、どのような協調戦略の学習を行なうのかを観察する。

2 販売ゲーム

ゲームは図1のように、複数の経営主体とそれらの相互作用によって構成される[佐藤他, 1994]。

2.1 市場

市場の概要

商品は1つだけであり、これをひとつの販売地域で販売する。景気変動はなく需要一定であって全体需要は2,400(個)である。したがって、もし4社で同じような販売力で競合するなら2,400の4分の1程度の販売量を期待できる。

販売員

自社が契約した販売員が自社製品を売る。契約可能な販売員の数は制限がある。販売員には売上に無関係に定額給料を一人1期あたり150万円支払う。販売員の採用と解雇は直ちに行うことができる。販売員がいなければ、売上はない。

商品価格

商品の価格は、15万円から1万円刻みに20万円までの、6種類から選ぶこととする。

仕入れ

仕入れ価格は商品1個あたり144,000円で一定とする。ある期に発注した量は次の期首にかならずきちんと納入される。

受注の決まり方

受注は、価格と販売員数と広告数で決まる。製品の市場価格が低いほど受注しやすい。維持している販売員数が多いほど、受注しやすい。販売員を増やすことへの受注への効果は、販売員数について逓減することなく、員数に比例した効果を持つ。新聞広告のページ数が多いほど受注に結び付きやすく、受注への効果は広告ページ数に比例する。やはり他社と競合し受注に結びつく。蓄積効果はない。広告は、1商品1ページ当り、10万円かかる。

当期の受注が、当期の販売可能数を上回った場合でも、受注残はない。しかも、のれんには傷はつかない。

2.2 企業

企業が各期に決定すべきことは次の2つの計画である。

- 1) 販売計画: 販売価格、契約する販売員数、広告ページ数、仕入れ量
- 2) 資金計画: 借入金(商品仕入、販売コスト、支払い利息にあてられる)

これらの計画から市場における企業間競争が具体的に決まり、各社の受注量が需要の取り合いの結果として決まる。

借入金は1期間借りることができ、次の期の期首に元金と利息を返済する。利率は3%で一定である。

商品の売上は手形で支払われ、次の期の期首に現金となるものとする。割引はしない。

以上で定義された販売ゲームは[佐藤他, 1994]と類似しているが、「販売員がゼロならば売上もゼロ」という設定になっている。直感的には今回のほうが学習しにくいと考えられる。

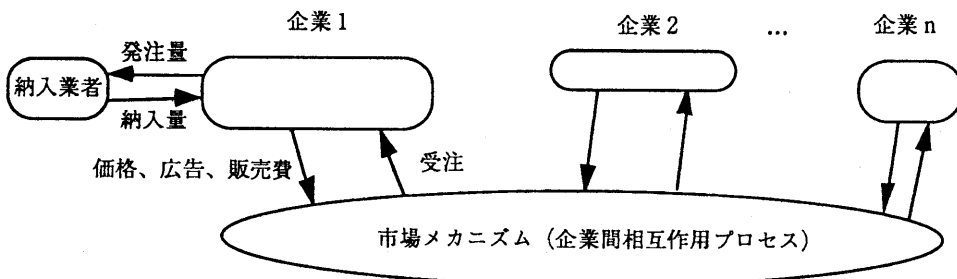


図1 ゲーミングの設定

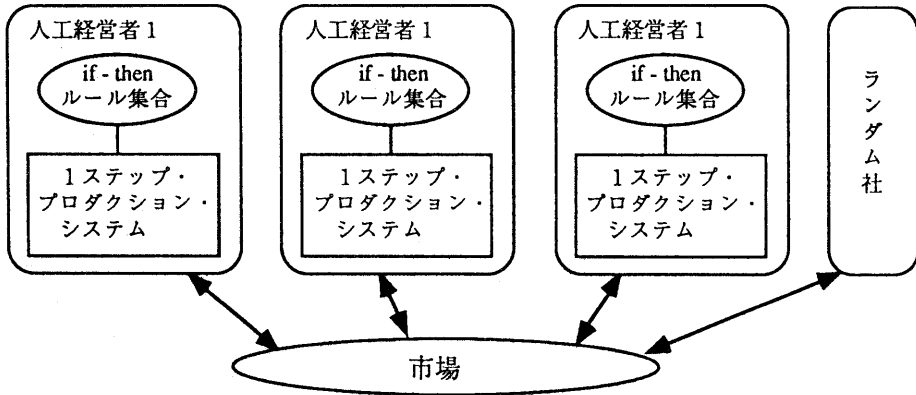


図2. 3社の学習とゲーミング

3 販売ゲームの人工経営者

人工経営者の構成は[佐藤他, 1994]と類似しているが、条件部を構成する属性が17個から4個減り13個となっている。

3-1 基本構成

人工経営者はif-then型ルールに基づく生産システムを使って人工経営者を作る。if-then型ルールは「もし今期のデータについてある条件が成立するならば、次期の販売計画の決定変数の値をある数値だけ増加する」という形で意思決定規則を表している。たとえば、「もし販売単価が16万円以下で利益が1300万円以下ならば、単価を1万円上げて、かつ仕入れを30個減らせ。」である。よい経営をするためにはよいルールの集合が必要となる。販売ゲームを繰り返す過程でルールをそろえることは、経営の仕方をルールとして表現することにほかならない。これはルール獲得や

ルール学習と呼ばれる。ルール獲得をGAを通じて「経験的に行なう」機械学習システムとして、クラシファイア・システム(classifier system)がある[樋口・北野(1993);計測と制御(1993)]。獲得されたルール集合にもとづく経営の結果は、貸借対照表や損益計算書の初期値に依存して、異なる経営結果をもたらす。ちょうど、チューリングマシンのプログラムが人工経営者のルール集合であり、チューリングマシンのテープ上の初期値とヘッド制御部の初期値が人工経営者にとってのB/SとP/L(貸借対照表と損益計算書)の初期値に対応する。チューリングマシンでも初期値が異なると、対応して異なる計算を行ない異なる結果が得られる。

3-2 ルール学習の仕組み=クラシファイア・システム

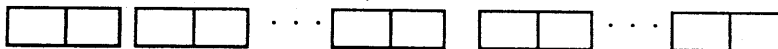
本研究のクラシファイア・システムは、[佐藤他, 1994]と同様の実数を直接扱えるものであり、fGAC

ひとつの属性



演算子 属性値

ひとつのルール (17個の属性からなる)



属性1

属性2

属性13

属性1

属性4

条件部 (13個の属性)

結論部 (4個の属性)

図3 fGACのデータ構造

(floating-point-representation-GA-based Classifier system)と呼んでいる。従来のクラシファイア・システムはGAによるルール制御、またそのルール自体の表現形式のためにバイナリ表現へのコーディングが必要であったが、fGACはこのコーディングを必要としない。実数表現したデータの遺伝的アルゴリズムは[Michalewicz, 1992]で紹介されている。

fGACのひとつのルールはクラシファイアと呼ばれる。「たとえば、現在の販売員数が302人より多い」というひとつの事実は、「>, 302」のように演算子と属性値の組で表わす。条件部(if部)と帰結部(then部)から成る。図3のように表現される。

条件部、帰結部ともに幾つかの属性から構成される。ひとつの属性は演算子と属性値の組から成る。属性値は実数値である。各ルールの条件部には前期の決定変数及びB/S,P/Lからの情報13属性を割り当てる。それは、商品価格、広告量、販売員数、仕入量、売上高、売上原価、一般管理費、支払利息、利益、期首現金、商品期首在庫、借入金、及びマーケット・シェアである。演算子は条件部と帰結部で異なる。条件部における属性の演算子は"#", ">", "<"の3種類、帰結部のそれは"+", "-", "noop(eration)"の3種類がある。条件部の"#は無視(don't care)記号であり、ゲームの状況とルールの条件部をマッチングする時に、その属性についてどのようなものであっても構わないことを意味する。したがって#に対応する属性はそのルールに存在しないことと同じ意味を持つ。演算子">"はゲームの状況の該当する属性値がそのルールのもつ属性値よりも大きいこと、"<"は小さいこ

属性番号	属性	下限	上限
0	商品価格(万円)	15	20
1	広告量(ページ)	0	140
2	販売員数(人)	0	500
3	仕入量(個)	0	6000
4	現金(万円)	-1000000	1000000
5	売上高(万円)	0	200000
6	商品(個)	0	10000
7	借入金(万円)	0	1000000
8	利益(万円)	-1000000	1000000
9	売上原価(万円)	0	144000
10	支払利息(万円)	0	20000
11	一般管理費(万円)	0	1000
12	マーケットシェア	0	1.0

表1 各属性値の範囲

とをそれぞれ意味し、この演算子と値のセットを用いて自社の状態(B/S,P/L)とルールのすべての属性間でのマッチングが図られる。

帰結部に割り当てる属性は商品価格、広告量、販売員数、仕入量の4つがある。

各属性値の範囲は表1の通りである。fGAC内で保持されるルール集合全体(クラシファイア・リストと呼ばれる)は図4のようになる。各ルールは、各ルールの一種の有効性を表した信頼度という数値と、そのルールがfGACによるゲーミング・シミュレーションで何回使用されたかをカウントする使用回数の2つの数値を持つ。このルール集合を遺伝的アルゴリズム操作によって学習させていくことになる。今回の実験ではルール数は200である。

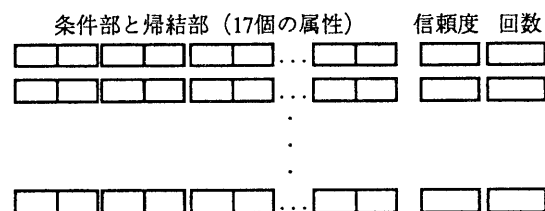


図4 クラシファイア・リスト

3-3 fGACの動作概要

fGACによるルール集合の学習(ルール獲得)は、図5のように行なった。この詳細については[佐藤他, 1994]を参照されたい。

「信頼度の割り当て計算」では、ゲームを1期だけ行った結果からそこで用いたひとつのルールの信頼度を変更する。ここでの評価項目に何を用いるか、さらにどのような結果が生じたときにどのくらいの報酬を与えるかが問題となるが、ここでは表5にまとめたような評価項目でそのルールに対する報酬を決定し、信頼度の操作を行った。

表2の通り、報酬は大別して成功したときの正の報酬、失敗したときの負の報酬の2通りの与え方で行う。良い評価を与えたルールの信頼度に関しては規定された数値をその信頼度の値に加えるが、良くない評価を与えたルールの信頼度に関してはそれから規定の数値を減じる。

3-4 ルール発見システム

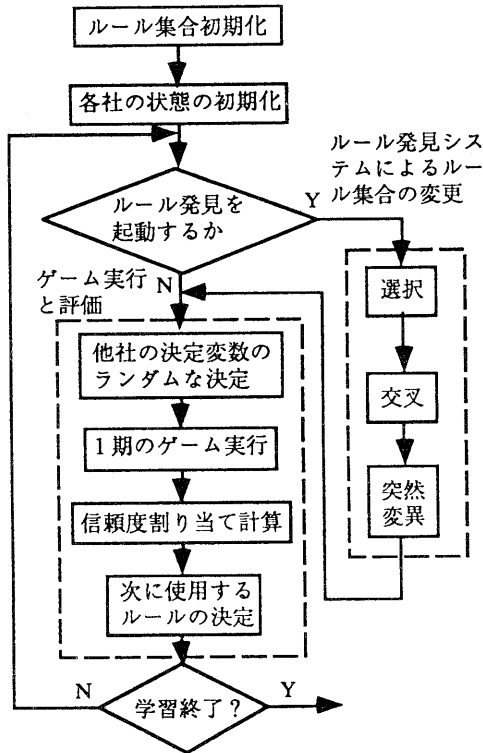


図5 fGACによるルール学習

本システムでは選択されたルールに対して、次の世代*t+1*における信頼度*C(t+1)*は報酬*R(t)*を素朴に加えた（負の場合には減じた）。この結果、始めに選ばれたルールがかなり影響するような、非常に脆いルール学習を行なうことになる。

ルールの帰結部の"+", "-"は属性値に示された値を今までの決定変数の値にそれぞれ加算、減算して新たな決定変数の値とする。

属性	条件	報酬
利益の変化	値が前期より増加	+20
	値が前期よりも減少	-10
利益-在庫	値が前期より増加	+20
	値が前期より減少	-10
借入金	前期より減少	+30
	前期より増加	-20
マーケットシェア	前期より増加	+10
	シェアゼロ(販売なし)	-20

表2 信頼度計算時の評価項目

ルール集合の更新をするためにfGA操作を行うルール発見システムが起動される。クラシファイア・リストを個体の集団とみなし、その中の個体であるクラシファイアに対して遺伝的操作を施す。このルール発見システムを何世代毎に起動させるかは予め与えておく。選択（淘汰）、交叉、突然変異という一連の遺伝的アルゴリズムの操作は以下のように行なう。

選択では各ルールの信頼度を適応度として、各ルールの信頼度の比率に等しい確率で新たな集団におけるルールをサンプリングして選び出す。選択を終えた集団内のルールに対して単純な一点交叉を行う。交叉率を予め決めておき、乱数を用いて各ルールへの操作の有無を決定する。交叉点の位置はルールの条件部、帰結部を問わず乱数を用いて決定され、それを起点に情報の交換がなされる(図5)。

なお、交叉が行われるときにはそのルールであるfGACに付随する信頼度、回数といった情報も染色体の一部として取り扱う。新たに生成された染色体であるfGACの信頼度*c₁'*、*c₂'*はその親染色体であるfGACの信頼度*c₁*、*c₂*を用いて以下のように定める。

$$c_1' = \text{rand} * c_1 + (1 - \text{rand}) * c_2$$

$$c_2' = \text{rand} * c_2 + (1 - \text{rand}) * c_1$$

ここで、randは[0,1]の一樣乱数である。子染色体の回数情報は新たに0にセットする。

突然変異の対象となるのは条件部と帰結部の属性演算子及び値である。

3-5 機械学習実験

GAにおける集団サイズ数(population size)にあたるクラシファイア・リストのルール数を200とした。fGACによるルール学習フェーズでは、ゲームの総繰り返し回数は100期であり、その中でfGAによるルール発見システムを10世代に1回起動させ、23期ごとには初期B/SとP/Lをランダムに規定範囲内で変更する。よってfGACによるルール学習フェーズでのルール発見システム起動回数は100/10で10回である。また、交叉率0.25、突然変異率0.01とする。

fGACによって獲得されるルール集合は、ゲーミングシミュレーションにおいて対戦相手3社が用いる乱数が異なると、かなり異なるものになる。乱数が異なるということは、初期値である初期B/SとP/Lが異なるということと、学習過程で行う4社

競合の販売ゲームでの価格付けや投入販売員数や広告数が違うということになり、その結果、学習されたルールが違う。さらに、その学習されたひとつのルール集合についても、ゲーミングを行うときの自社のB/SとP/Lに依存して、ゲームのふるまいが違う、つまり使用されるルールとその結果の経営成績が多様になる[10]。うまく行く場合もあれば、まずい場合もある。需要は充分あるので、150人ぐらいの販売員を雇い、4000個ぐらいの商品を仕入れ、高価格を付け、その後それらを変化させなければ、利益を出せる。うまく学習した場合のルールは、販売員が0だと売上がゼロであることも結果として学習していると見させる。

4 協調戦略の創発

前節までの学習シミュレーションは、学習する人工経営者はただ1社であった。今度は、3社が学習するものとする(図2)。1社の場合と同様に、まず学習フェーズで3社のルール学習が同時に行なわれる。これはfGACの図で「他社の決定変数のランダムな決定」のところ、学習する3社はもっとも適したルールをひとつずつ選んで適用し、他のランダム社は、ランダムな決定をするということである。

この場合に、3社が結託して高価格路線をしけばランダム社に対して非常に有利なビジネスを行

なうことができるが、そのような協調戦略を発見できるかということが注目点である。

結果として、3人の人工経営者は高価格路線の(異なる)ルール集合を発見する。しかし、3人のそれら学習済み人工経営者と1人のランダム経営者とのゲームの結果は、B/S、P/Lの初期値と、ルール集合に依存するために、うまく経営するもの、そうでないものに分かれる。

参考文献

- [1] Axelrod, R.: The evolution of strategies in the iterated prisoner's dilemma, in Davis, L. (editor) "Genetic Algorithms and Simulated Annealing," pp32/41, Morgan Kaufmann Publishers, 1987. [2] 樋口哲也、北野宏明: 遺伝的アルゴリズムとその応用、情報処理、vol. 34, No.7, pp871/883, 情報処理学会、1993. [3] Michalewicz, Z.: Genetic Algorithms + Data Structures = Evolution Programs, Springer, 1992. [4] 佐藤亮: 遺伝的アルゴリズムはゲーミングの解を見つけるか、経営情報学会1993年秋季大会講演要旨、pp143-148, 1993. [5] 佐藤、佐藤、松浦: 人工経営者は販売ゲームの解を見つけるか、Inst. of Socio-Economic Planning Discussion Paper No. 572, Univ. of Tsukuba, 1994. [6] 佐藤亮: 資本主義乱立のAManシミュレーション、経営情報学会1994年春季大会講演要旨、1994. [7] 竹内勝: 遺伝的アルゴリズムによる機械学習、計測と制御、32-1, pp24-31, 1993. [8] 松尾和洋: ゲーム型生態系における自己組織化、計測と制御、Vol. 29, No.10, pp41-46, 1990. [9] 計測と制御、vol.32, No. 1, 特集 遺伝的アルゴリズム、計測自動制御学会、1993. [10] 佐藤亮: 人工経営者へのクワリアイトによるアプローチ、システム情報関連シンポジウム予稿集、計測自動制御学会、1995.

