推薦論文

部分的完全調査モンテカルロ法の提案と TAC エージェントの意思決定への応用

川村 秀憲[†] 小野寺 将輝[††] 大内 東[†]

コンピュータゲームや電子商取引などにおけるエージェントの意思決定を設計する際，状況によっては，ある意思決定に対して生起する結果が一意に決定されないという不確実性を含む場合があり，生起する結果の確率を考慮に入れた意思決定の手法が必要となることがある．そのような際，生起する結果の評価値の期待値を用いる方法が考えられるが，計算コストに対する制約などにより実際に期待値の真値を計算することが困難な場合は，モンテカルロ法を用いて推定することが一般的である．本論文では，生起確率が高い結果の列挙が容易である場合にサンプル数を増加させることなく期待値の推定精度を上げる部分的完全調査モンテカルロ法を提案し，実験により提案手法の有効性を確認した．また応用例として，TAC CLASSIC という多人数参加の取引ゲームのエージェント設計へ適用した．実験の結果，提案手法を組み込んだ意思決定が不確実性を含むゲームにおいて有効であることが示された．

Proposal of Partially Exhaustive Investigation Monte Carlo Method and Application to Decision-making of the TAC Agent

HIDENORI KAWAMURA,[†] MASAKI ONODERA[††] and AZUMA OHUCHI[†]

To design an autonomous-trading software agent for some kinds of game and electric commerce, the agent often has to cope with the situation that result of decision-making includes stochastic phenomena. In such case, it is necessary for the agent to consider expected evaluation value for possible results according to the probability of results, and Monte Carlo Method is generally used to calculate the expected evaluation value for such situation. In this paper, we suppose the situation that it is easy to enumerate high-probability results at low computation cost, and propose an improved version of Monte Carlo Method, i.e., Partially EXhaustive Investigation (PEXI) Monte Carlo Method. Moreover, we show the effectiveness of the proposed method by some computer experiments including the designing of TAC CLASSIC trading agent.

1. はじめに

コンピュータゲームや電子商取引などの分野において，ソフトウェアによる意思決定の研究が行われている．そこでは，将棋や囲碁，ダブルオークションのように確率事象を含まない決定論的な状況や，麻雀，トランプやさいころを用いたゲーム，確率的価格変動モデルを含むオークションなど，確率事象を含むものまで幅広い状況での意思決定に関して研究がなされている．ここでは，後者のような状況に焦点を当てる．

ここでの確率事象を含む意思決定とは，ある状況において選択可能な選択肢が複数存在し，さらにそれぞれの選択肢において複数の異なる結果が確率的に生起する状況を想定する．また，確率事象を含まないゲームにおいても，状況が確率的にしか判断つかない場合や全探索によって有効な手を発見することが困難な場合などに確率的なモデルを導入することで都合が良い場合がある[1),2)]．このような状況での意思決定では，いかに確率を考慮に入れた評価関数を設計するかが重要である．

† 北海道大学大学院情報科学研究科
Graduate School of Information Science and Technology, Hokkaido University
†† 株式会社 NTT ドコモ北海道
NTT DoCoMo Hookaido, Inc.

本論文の内容は 2004 年 5 月の情報処理北海道シンポジウム 2004 にて報告され，情報処理北海道シンポジウム 2004 プログラム委員により情報処理学会論文誌への掲載が推薦された論文である．

ここで，ある選択肢に対して起こりうる結果の生起確率が既知である場合について考える．たとえば，ルーレットではどの数字に賭けるのかを選択した後に起こりうる結果を確率的に知ることができるし，ポーカーなどでは交換する枚数を選択した後に配布される可能性のあるカードの確率を計算することができる．このような場合，各結果に対する静的評価関数を設計し，選択肢に対して生起する可能性のある結果の評価値の期待値を求めることで意思決定を行うことが一般的であるが，状況によってはある選択肢に対して考慮すべき結果が膨大に存在し，期待値の真値を計算することが困難な場合がある．

そのような場合，期待値の推定にモンテカルロ法（以下，MC 法）[1),3)] を用いることができるが，MC 法を意思決定に用いる際には計算コストの観点から，より少数のサンプルで可能な限り精度を向上させたいという要求が生じる．ここでもし，起こりうる結果が複数の独立な確率事象の組合せで定義され，かつそれぞれの事象での確率の分布に偏りが大きい場合，確率の大きい事象から順に確率の積を計算して列挙することによって，全事象を考慮しなくても生起確率の高い結果を列挙可能な場合がある．この特徴に基づいて期待値への影響が大きい結果を有効利用することによって，それほど計算コストを増加させることなく，同じサンプル数で MC 法の期待値推定の精度を向上させることができる余地がある．

本論文では，上記のような状況のもと，同じサンプル数による MC 法を用いた期待値の推定の精度を向上させる部分的完全調査 MC 法（Partially EXaustive Investigation MC 法，以下 PEXI-MC 法）を提案する．まず，数式の展開により精度向上の可能性を示したのち，予備実験によって PEXI-MC 法の有効性を確認する．さらに，応用の一例として，確率的価格変動モデルを含む TAC CLASSIC という複数エージェント参加型のゲームへの適用を行い，提案手法を用いた意思決定の有効性に関する考察を行う．

2. 部分的完全調査モンテカルロ法の提案

2.1 モデル化

ある意思決定の場面においていくつかの選択肢が存在し，それぞれの選択肢に対して生起する結果が有限で互いに独立な確率事象の組合せとして定義できる状況を想定する．各事象の生起確率と，それぞれの結果の評価値は既知であるとする．このような状況で意思決定する際に，起こりうる可能性がある結果の評価値の期待値が最大の選択肢を選択するために，その期待値を計算することを考える．

まず，選択可能な選択肢の集合を S とする．

$$S = \{s_i \mid i = 1, 2, \cdots, N_S\} \tag{1}$$

ある選択肢 $s \in S$ に対して確率的に生起する結果の事象集合 X^s を以下のように定義する．

$$X^s = \{x_i = (y_1, \ldots, y_N) \mid \tag{2}$$

$$y_1 \in Y_1^s, \ldots, y_N \in Y_N^s\} \tag{3}$$

$$Y_j^s = \{y_{j1}, \ldots, y_{jN_j}\}, j = 1, 2, \ldots, N \tag{4}$$

このとき，x_i の生起確率を $p(x_i)$，y_{jk} の生起確率を $p(y_{jk})$ とすると，$p(x_i)$ は $p(y_{jk})$ の積として．

$$p(x_i) = p(y_1) \times \cdots \times p(y_N) \tag{5}$$

のように計算される．ただし，

$$\sum_{x \in X^s} p(x) = \sum_{y \in Y_j^s} p(y) = 1 \tag{6}$$

である．

起こりうる結果 x_i の評価値を $e(x_i)$ とすると，選択肢 s の評価値の期待値 E は，

$$E = \sum_{x \in X^s} p(x) \cdot e(x) \tag{7}$$

と計算される．

意思決定ではこの期待値が最大となる選択肢 s を選択したい．ここで，期待値の真値が現実的な計算コストで計算できればよいが，選択肢が多数あったり，生起する結果の組合せが膨大だったりする場合は真値を計算することが容易ではない．ここでは，そのような場合にいかに限られた計算コストで精度良く期待値を推定するかについて議論を行う．

2.2 生起確率が高い結果の列挙

前述のように，ある結果が複数の独立な確率事象の組合せとして定義される場合，その生起確率を求める過程では確率の積を計算することになる．このとき，結果に対する生起確率にある閾値を設定し，積の計算過程において閾値を下回った時点でそれ以降の計算をうち切る単純な枝狩りによって，生起確率が閾値以上のものを列挙するアルゴリズムを容易に設計できる．ここではそれに基づいたアルゴリズムを利用する．

図 1 のアルゴリズムでは，閾値 *threshold* を設定して *enumerate*1() を呼び出すことで，閾値以上の生起確率を持つ結果を列挙することができる．アルゴリズムでは積の計算過程で閾値未満だったときにそれ以上の積の計算をうち切ることで枝狩りを行う．

閾値の設定に関しては確率の偏りに応じて調整が必要であり，効率の良い設定は問題ごとに異なってくると思われる．実際の問題に応用する際には列挙の過程で自動調整するなどのアイデアが考えられるが，ここ

```
enumerate1() {
    enumerate2(1.0, 1);
    保存した結果を出力;
}
enumerate2(float prob, int j){
    for(int k = 1; k < N_j; k++) {
        float prob* = prob · p(y_jk);
        if (prob* ≥ threshold){
            if (j == N) 調査した結果 x_i^s を保存;
            else enumerate2(prob*, j + 1);
        }
    }
}
```

図 1 列挙のアルゴリズム
Fig. 1 Enumerating algorithm.

では閾値の調整は実験的な取扱いとする．また，生起確率の偏りの大きさを考慮に入れて計算順序をソートすることでよりアルゴリズムの効率を追求することも可能であるが，簡単化のためにソーティングの手順はアルゴリズムに含めないもとする．

2.3 部分的完全調査モンテカルロ法

図 1 のアルゴリズムに従って生起確率の高い結果を列挙することが可能であるという特徴を利用し，列挙した部分の完全調査を組み合わせることで MC 法の精度を向上させる PEXI-MC 法を提案する．MC 法の精度を評価する一指標として，推定結果の分散を用いることができる．つまり分散が小さいほど推定精度は高いと考えることができる．そこで，提案手法を用いた場合の分散と，基準となる加重サンプリング初等的 MC 法を用いた場合の分散との比較により，提案手法の評価を行う．

まず加重サンプリング初等的 MC 法[4] を用いた場合の X^s に関する期待値の推定について説明する．発生させる総サンプル数を M とし，$p(x_i)$ に従って X^s から発生させた M 個のサンプルを $\xi_1, \xi_2 \cdots, \xi_M \in X^s$ とすると，得られる期待値の推定値 E_1 とその分散 $varE_1$ は，

$$E_1 = \frac{1}{M}\sum_{i=1}^{M} e(\xi_i) \tag{8}$$

$$varE_1 = \frac{1}{M} var\ e(\xi) \tag{9}$$

となる．

次に，PEXI-MC 法を用いた場合の X^s に関する期待値の推定について説明する．まず，図 1 のアルゴリズムに従って X^s から列挙された閾値以上の生起確率を持つ事象集合を X_1^s とし，それに含まれない事象集合を X_2^s とする．X_1^s に含まれる要素数を $|X_1^s|$ とし，それぞれの事象の調査に使用するサンプル数を M_1, M_2 とすると，$M = M_1 + M_2$ であり，$M_1 = |X_1^s|$ である．

まず，手順 1 として，X_1^s に関して完全調査を行う．X_1^s による期待値の推定への寄与値を D_1 とすると，X_1^s についてはすべての要素を考慮するので，

$$D_1 = \sum_{x \in X_1^s} e(x) \cdot p(x) \tag{10}$$

となる．

手順 2 として，X_2^s による寄与値を加重サンプリング初等的 MC 法で計算する．X_2^s に含まれる要素 $x \in X_2^s$ の生起確率を $p_2(x)$ とすると，

$$p_2(x) = p(x) / \sum_{x' \in X_2^s} p(x') \tag{11}$$

となる．$p_2(x)$ に従って発生させた $M_2 = (M - M_1)$ 個のサンプルを $\xi_1^*, \xi_2^*, \cdots, \xi_{M_2}^*$ とすると，X_2^s による寄与値 D_2 は，

$$D_2 = \sum_{x \in X_2^s} p(x) \cdot \frac{1}{M_2} \sum_{i=1}^{M_2} e(\xi_i^*) \tag{12}$$

$$= \left(1 - \sum_{x \in X_1^s} p(x)\right) \cdot \frac{1}{M_2} \sum_{i=1}^{M_2} e(\xi_i) \tag{13}$$

となる．この式によって，X_2^s の全要素を調査することなく寄与値を計算することができる．

手順 3 として，PEXI-MC 法による期待値の推定値は以下のように計算される．得られる期待値の推定値を E_2，その分散を $varE_2$ とすると，

$$E_2 = D_1 + D_2 \tag{14}$$

$$varE_2 = \left(\sum_{x \in X_2^s} p(x)\right)^2 \frac{1}{M_2} var\ e(\xi^*) \tag{15}$$

$$= \left(1 - \sum_{x \in X_1^s} p(x)\right)^2 \frac{1}{M_2} var\ e(\xi^*) \tag{16}$$

である．分散が小さいほど精度が高いという考え方より，

$$varE_1 > varE_2 \tag{17}$$

を満たす場合には，提案手法による精度の向上が可能であると考えられる．そこで，いくつかの予備実験を行い手法による精度の違いを比較する．

2.4 予備実験の設定

PEXI-MC 法と加重サンプリング初等的 MC 法との比較によって提案手法の有効性に関する予備実験を

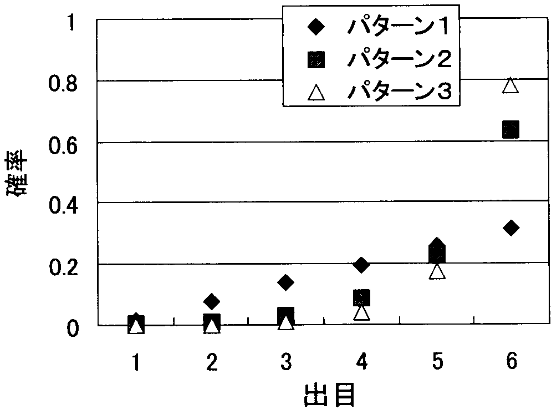


図 2 各目の生起確率（3 パターン）
Fig. 2 Probability of each sample.

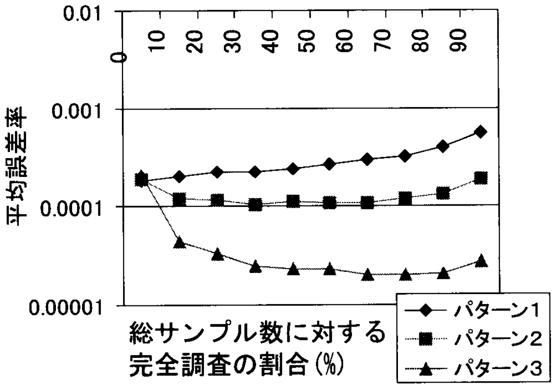


図 3 予備実験 1 における平均誤差率
Fig. 3 Average error in preliminary experiment 1.

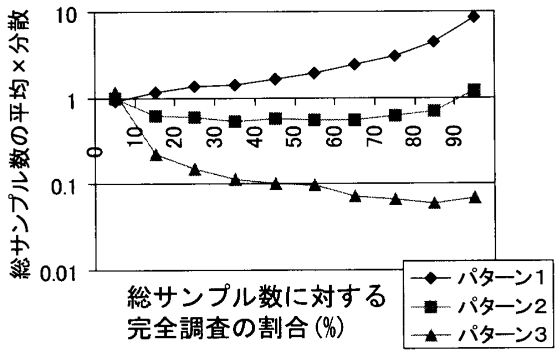


図 4 予備実験 1 における総サンプル数の平均 × 分散
Fig. 4 Number of total samples times deviation of expected value in preliminary experiment 1.

行う．実験設定として，出目の確率に偏りのある 8 個のさいころを同時にふる場合を考える．すなわち，それぞれのさいころの出目が Y_j^s に対応し，8 個のさいころの出目の組合せ結果が X^s に対応する．実験においては，出目の組合せごとにランダムな評価値を事前に設定しておき，各手法で期待値の推定を行う．この実験では期待値の推定のみ扱い，選択肢の選択については取り扱わない．

PEXI-MC 法の性質上，完全調査される部分の確率によって推定精度が変わってくると思われるので，実験で用いるさいころの出目の生起確率は，偏りが異なる**図 2** のような 3 種類のパターンを用意した．8 個のさいころとそれぞれ 6 つの出目であるので，$|X^s| = 1{,}679{,}616$ である．出目の組合せに対する評価値は，平均 100，分散 1 で発生させた正規乱数を割り当てることとした．少ないサンプル数で推定することを目的とするため，発生させる総サンプル数 M は全組合せ数の約 0.1% として $M = 1{,}680$ 個とした．

PEXI-MC 法の実装にあたって，図 1 のアルゴリズムのみで X_1^s を決定すると閾値と生起確率によって完全調査に含まれる数が変わってしまう．予備実験では総サンプル数にしめる完全調査の割合の影響を調べて提案手法の有効性の検討を容易にするため，十分低い閾値をもってアルゴリズムで列挙したのち，確率が上位のものから一定数を指定して X_1^s とすることで完全調査の個数を指定する．また，X_2^s に含まれるサンプルのみをサンプリングするためには事前に $p(x)$ に基づいて $p_2(x)$ を計算しておく必要があるが，実装の簡略化のために $p(x)$ に従ってサンプルを発生させ，もしそのサンプルが X_1^s に含まれた場合は棄却して再度サンプリングを行うことで規定のサンプル数を確保するものとし，確率の再計算のコストを省くものとする．

予備実験 1 では，ある評価値のパターンに対して完全調査を行う個数 M_1 を総サンプル数 M の 0～90% まで 10% 刻みで変化させ，それぞれ 1,000 試行の推定を行い，|(推定値) − (真値)| ÷ (真値) で得られる誤差率の平均を比較する．予備実験 2 では，評価値のパターンを 10 種類用意し，各 100 試行ずつ，合計 1,000 試行を行い誤差率の平均を比較する．また予備実験 2 では，完全調査を行う個数 M_1 を総サンプル数 M の 5% 刻みで変化させることとする．

2.5 予備実験の結果・考察

図 3 は，実験 1 における 1,000 試行の誤差率の平均である．完全調査の割合が 0 である一番左のプロットが加重サンプリング初等的 MC 法による結果に相当する．結果を見てみると，生起確率の偏りが小さいパターン 1 では PEXI-MC 法によって誤差率が増加しているが，パターン 2，3 のように生起確率の偏りが大きい場合には初等的 MC 法よりも PEXI-MC 法の方が精度が高くなる部分が存在することが確認できる．

予備実験における PEXI-MC 法の実装では使用せずに破棄してしまった無駄なサンプルが発生しているので，破棄したサンプルに関するコストも考慮する必要があると考えられる．そこで，実際に発生させた総サンプル数と推定値の分散の積を**図 4** に示す．グラ

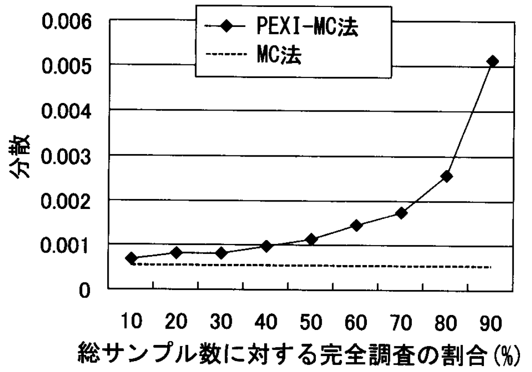


図 5 予備実験 1（パターン 1）における分散の変化
Fig. 5 Change of deviation of pattern 1 in preliminary experiment 1.

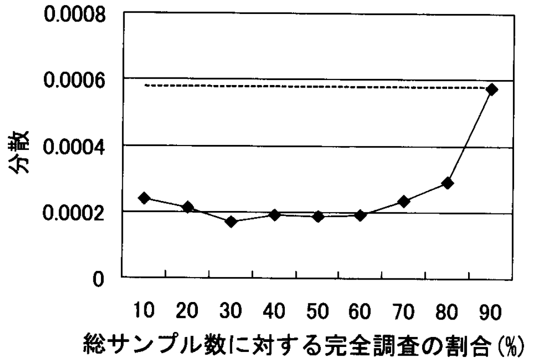


図 6 予備実験 1（パターン 2）における分散の変化
Fig. 6 Change of deviation of pattern 2 in preliminary experiment 1.

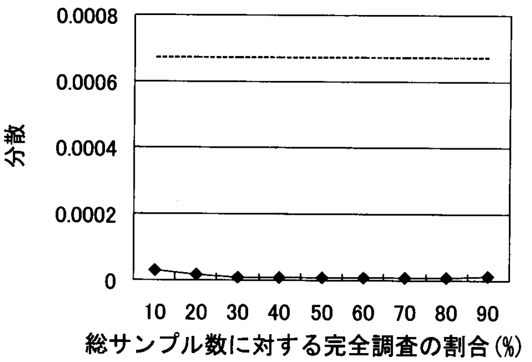


図 7 予備実験 1（パターン 3）における分散の変化
Fig. 7 Change of deviation of pattern 3 in preliminary experiment 1.

フから，破棄したサンプルのコストを考慮しても初等的 MC 法よりも有効な場合が存在することが確認できる．

図 5，**図 6**，**図 7** は，実験 1 の 1,000 試行において完全調査の割合を変化させた場合に PEXI-MC 法によって得られた推定値の分散を表す．パターン 1 ではすべての割合において初等的 MC 法による分散が PEXI-MC 法による分散を下回っている．それに対してパターン 2 と 3 では PEXI-MC 法の方が分散が小

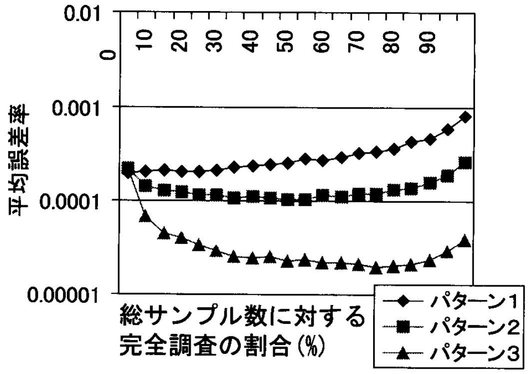


図 8 予備実験 2 における平均誤差率
Fig. 8 Average error in preliminary experiment 2.

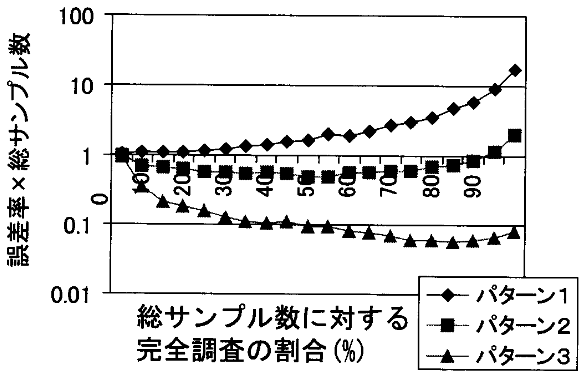


図 9 予備実験 2 における総サンプル数の平均 × 分散
Fig. 9 Number of total samples times deviation of expected value in preliminary experiment 2.

さくなっている．これらより式 (17) を満たす場合には，実際に精度の向上を図ることができることが確認できる．

図 8，**図 9** は，予備実験 2 における誤差率の平均，および実際に発生させたサンプルの総数と分散との積を示したグラフである．結果より，各結果に割り当てられる評価値が変化しても，実験 1 と同様の傾向が見られることが確認できる．

3. TAC CLASSIC

Trading Agent Competition（以下，TAC）とは，様々な分野における自動的に取引を行うエージェントに関する研究を行っている国際フォーラムである[5)~8)]．ここでは TAC の研究シナリオの 1 つである TAC CLASSIC を取り上げる．

TAC CLASSIC では，参加者として 8 人のエージェントがゲームに参加する．1 ゲームは実時間の 12 分間で行われ，それぞれのエージェントは各々に割り当てられた 8 人の顧客の好みに合わせて複数のオークションからチケットを入手し，各顧客に 5 日以内の旅行パックを提供する．顧客の好みを満たすほど高いス

コアが与えられるが，チケットの入手は確率的要素を持つものを含む複数種のオークションで行わなければならない．また，相手のエージェントの情報は入札情報などから限定的にしか判断できない．

エージェントの意思決定ではどういう入札を行うかを決定することになるが，オークションの結果によって入手できるチケットの数が変わってしまうので，入手できなかったチケットがある場合のリスクを含めて意思決定を行う必要がある．また，意思決定に用いることができる時間も制限されているので，限られた時間でできるだけ正確に期待値を最大化する意思決定が求められる．これらの特徴より，TAC CLASSIC は PEXI-MC 法を組み込んだ意思決定の有効性を検証するのに妥当な応用例であると考えられる．

旅行パックを構成する商品は，往復のフライト，ホテル，娯楽の 3 種類のチケットであり，概要は下記のとおりである．

フライトチケット 1〜4 日目の往路，2〜4 日目の復路のチケットが存在し，Single Seller Auction で売り出される．売り切れはないが，一度購入したチケットを売ることはできない．価格は確率的価格変動モデルに基づいて 24〜32 秒ごとに更新される．多くの場合，時間が経つにつれて価格は上昇する傾向を持つ．

ホテルチケット 1〜4 日目について TT，SS の 2 種類のホテルが存在し，TT は SS よりも良いホテルである．顧客は旅行中にホテルを変更することはできない．ホテルごとに 1 日 16 部屋が用意されている．チケットごとに Ascending Multi-unit Auction で取引され，エージェントは希望買い価格を入札する．市場は，4 分から 11 分の間に毎分 1 つずつランダムに終了し，終了時に 16 番目に高い入札価格を落札価格として上位 16 位までが落札できる．一度行った入札は取り消せず，一度入手したチケットを売ることもできない．

娯楽チケット 1 日目から 4 日目の各日にワニレスリング，遊園地，博物館のチケットがそれぞれ 8 枚ずつ存在し，顧客は出発日以外に 1 日 1 枚だけ使用可能である．ゲーム開始時に各エージェントへ一定枚数がランダムに配布される．エージェントは任意の価格で Continuous Double Auction に売買入札をすることができるので，価格は需要と供給のバランスによって決定される．

各顧客は，旅行パックに対して以下のような好みを持つ．

- 希望到着日（PA），希望出発日（PD）
- TT に泊まった場合のボーナス（HB）
- 各娯楽チケットに対するボーナス（AW，AP，MU）

ゲーム終了時の顧客の効用 U は以下のように計算される．

$$
\begin{aligned}
U &= 1000 - t\_penalty + h\_bonus + e\_bonus \\
t\_penalty &= (|AA - PA| + |AD - PD|) \times 100 \\
h\_bonus &= TT? \times HB \\
e\_bonus &= AW? \times AW + AP? \times AP \\
&\quad + MU? \times MU
\end{aligned}
$$

ここで，AA，AD は実際の到着日と出発日を表す．また，$TT?$ はその客が TT へ宿泊したかどうか，$AW?$，$AP?$，$MU?$ は各娯楽チケットが旅行パックに含まれていたかどうかを $\{0,1\}$ で表す．最終的なエージェントのスコアは，8 人の客の効用の総和からチケット入手にかかった総コストを引いて計算される．

4. エージェントの意思決定設計

TAC CLASSIC において，PEXI-MC 法を用いた PEXI-MC-Agent の意思決定を設計し，その有効性を検討する．ここでは比較対象として，加重サンプリング初等的 MC 法を用いた MC-Agent と期待値を考慮しない MAX-Agent も準備する．エージェントの意思決定においては，ゲームの最中にそれぞれのオークションにどのような入札を行うかを決定する必要があるが，そのためにはある入札に関して起こりうる結果の確率と落札価格の予測値を用いてスコアの期待値を計算し，それが最大となる入札を行う必要がある．ここでは意思決定に用いる落札価格の予測，GA による入札の探索，その適応度の順に説明する．

4.1 落札価格の予想

フライトチケットの落札価格は他のエージェントによる影響はなく一方的に提示されること，および娯楽チケットは Continuous Double Auction によって各時点での需給バランスを考慮した価格が提示されていることから，この 2 つのチケットに関しては簡易的に現在価格をそのまま予想落札価格として利用する．

一方，ホテルチケットではホテルオークションの特徴から落札価格が他のエージェントの意思決定や顧客の好みによって複雑に変化するので，単純な価格予測は困難である．よって，ホテルチケットの入札には TAC2003 Qualify Round の 590 ゲームのデータをもとに事前に落札価格を学習させたニューラルネットワーク（NN）を用いることとした．NN は入力層 3，中間層 8，出力層 1 の 3 層構成で，2 分前，1 分前，現

顧客1 ···· 顧客8 顧客の優先順

$$\boxed{t^1 e_1^1 e_2^1 e_3^1 e_4^1} \cdots \boxed{t^8 e_1^8 e_2^8 e_3^8 e_4^8} \boxed{r^1 r^2 r^3 r^4 r^5 r^6 r^7 r^8}$$

顧客iの旅程 $t^i = \begin{cases} -1 & \text{割り当て無し} \\ 0 & \text{日程:1〜2日 ホテル}:SS \\ 1 & \text{日程:1〜3日 ホテル}:SS \\ \vdots & \vdots \\ 19 & \text{日程:4〜5日 ホテル}:TT \end{cases}$

顧客iのj日目の娯楽割り当て $e_j^i = \begin{cases} 0 & \text{割り当て無し} \\ 1 & \text{ワニレスリング} \\ 2 & \text{遊園地} \\ 3 & \text{博物館} \end{cases}$

顧客iの優先順位 $r^i \in \{1,2,\ldots,8\}$

図 10 GA における遺伝子型

Fig. 10 Genotype of Genetic Algorithm.

在の価格を入力として 1 分後の落札価格を予想として出力する．重みは出力の二乗誤差が最小になるように GA で調整した．

4.2 入札の決定

エージェントの意思決定では，どのチケットにどの価格で入札を行うかを決める必要がある．まずどのチケットに入札を行うかの決定についてであるが，期待値の推定値をもとに探索しなければならないこと，探索のコストを調整しやすいこと，複数の解を同時に探索できることなど，この問題に望まれる特徴を考慮して GA を利用することとした．

GA の設計は TAC 参加チームの PaininNEC[7] を参考にした．成立を目指す 8 人の顧客の旅行パックを遺伝子型とし，すべてのパック成立に不足しているチケットに関して入札を行う．入札価格に関しては，フライトチケットは予想落札価格をそのまま入札し，ホテルチケットは予想落札価格に一定幅のマージンを上乗せして入札を行う．また，娯楽チケットについては予想落札価格をもとに，売り入札は時間とともに減少する修正値を，買い入札については増加する修正値を加えたものを入札価格とする．これらの入札に基づいて入札を行った際のスコアの推定期待値を適応度として探索する．

ゲーム中では，GA による探索と入札の決定はゲーム開始時と 3〜11 分の間の毎分 1 回，計 10 回行う．設計した遺伝子型は**図 10** のとおりであり，1 人の顧客の旅行パックを表す部分 8 人分にチケットの優先順位を付け加えたもので表される．チケットの優先順位とは，複数必要になった同じチケットが全部入手できない場合にどの顧客を優先するかを表す．遺伝的操作

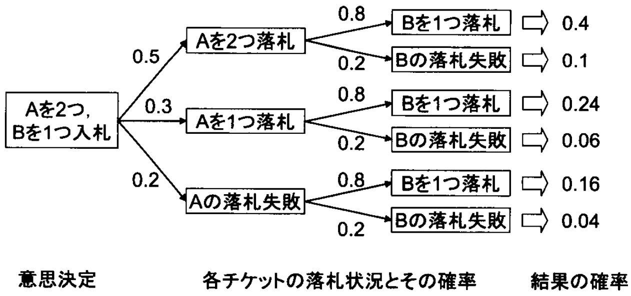


図 11 入札と結果，および生起確率

Fig. 11 Probabilities of combinations of bid and result.

としては，エリート保存戦略を併用したトーナメント選択，1 点交叉，突然変異を使用する．パラメータの設定は，探索に利用できる時間を考慮して集団数 200，世代数 100，交叉確率 0.9，突然変異確率 0.05 とする．

ここで，GA の各個体の適応度は期待値の推定値を利用するわけであるが，時間的制約から荒い推定しか行うことができないので，最終世代の最良個体をそのまま解とするのではなく，各世代での最良個体の期待値をより高い精度で再推定して選定し直すものとした．具体的には，各個体の期待値推定に用いるサンプル数を 100 から 1,000 に増やして再推定を行う．

4.3 適応度の計算

4.3.1 MC-Agent

MC-Agent では，入札によって得られるチケットの組合せの生起確率をもとに，起こりうる結果のスコアの期待値を加重サンプリング初等的 MC 法によって推定する．すなわち，チケットの組合せそれぞれの生起確率を用いて，その生起確率に従って結果をサンプリングしてその平均を計算することで推定値を計算することができる（**図 11**）．

このとき，ある入札に対して起こりうる結果の生起確率が必要であるが，各エージェントが判断できる状況は限定的であり，かつオークションの結果には毎回変動する確率的要素も含まれるため，ゲームのルールや実行時の状況から推定することは困難である．そこでここでは，予備実験の段階で後述する 7 つの Dummy-Agent を対象に 800 ゲームを行った際に得られた結果から計算された生起確率を用いるものとする．具体的には，参加エージェントの入札をランダムに決定し，チケットの種類に応じて入札数と実際に得られた数のデータベースを作成し，そのデータベースを利用して図 11 のように複数の入札に対する結果の生起確率を計算する．

また，MC-Agent で期待値の推定に利用するサンプル数は，実際に意思決定に利用できる時間を考慮して実験的に 100 とした．

4.3.2 PIXI-MC-Agent

PIXI-MC-Agent では，MC-Agent で加重サンプリング初等的 MC 法を用いて推定していた期待値を，PIXI-MC 法による推定に置き換えたものである．生起確率が高い結果については，図 1 のアルゴリズムに基づいて列挙を行う．列挙の際の閾値の設定に関しては，予備実験によって得られた生起確率の値の分布とアルゴリズムの実行時間を考慮して，10〜20 程度の結果が完全調査される 0.01 とした．この設定の妥当性については，後述する実験において検証する．

サンプリングに使用できるサンプル数は MC-Agent と同じ 100 とした．サンプルの抽出については，完全調査するサンプルを確定した後に全事象からランダムに加重サンプリング初等的 MC 法で使用するサンプルを作成し，そこから重複するサンプルを破棄する．その際，合計 100 サンプルに達するまでサンプリングを行うと計算コストが一定しないため，破棄されたサンプルについては補充しないものとした．

4.3.3 MAX-Agent

MAX-Agent の適応度は期待値の推定は行わず，すべて入札どおりのチケットが得られた場合のスコアをそのまま適応度として採用する．たとえば TAC2002 に参加した CHUK や PAININ NEC，ZEPP など多くのチームがこのような評価関数に基づいて入札を決定する[7]．

5. 実　　験

5.1 Dummy-Agent の導入

実験を行う際の対戦エージェントとして，TAC からサンプルとして提供される Dummy-Agent[8] に多少の変更を加えたものを用いた．Dummy-Agent は顧客の好みを重視して入札を行う．ホテルは HB が 70 以上なら TT，それ以外は SS とし，200 ドルから開始して以降は 50 ドルずつ競り上げる．娯楽チケットは，事前に決めた関数によって決定される価格で入札し，ボーナスの高い順に日程に含めるだけ含む．フライトチケットはすべて即購入する．

5.2 実　験　1

まず，TAC CLASSIC に PEXI-MC 法が有効かどうか検討するため，PEXI-MC-Agent と Dummy-Agent7 体で 1 ゲーム行い，その際に GA で生成した 20,000 個 ×10 回 =200,000 個の個体を例題として誤差率の平均を検証した．PEXI-MC 法の適用にあたって，実際のゲーム中では完全調査の個数を直接指定することは困難だが，ここでは事前に生起確率に従って結果をソートしたデータベースに基づいて完全調査の個数を指定した．それぞれの例題について，加重サンプリング初等的 MC 法，および完全調査を行う個数を 0〜90 個まで 10 個刻みで変化させた PEXI-MC 法の 2 種類の手法によって期待値を推定し，真値と比較した．

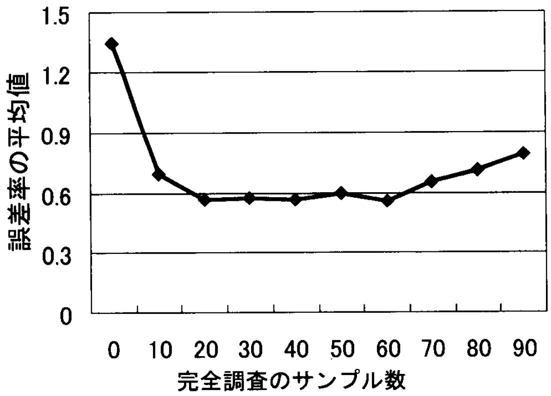


図 12 200,000 サンプルに関する誤差率の平均
Fig. 12 Average error in 200,000 samples.

図 12 は各例題に対して期待値を推定した際の誤差率の平均値である．横軸は完全調査のサンプル数，縦軸が誤差率の平均値である．完全調査のサンプル数が 0 の場合が加重サンプリング初等的 MC 法を用いた場合の結果にあたる．PEXI-MC 法によって誤差率が減少しているので，期待値の推定への有効性が確認された．また，グラフは省略するが，推定値の分散に関しても図 12 に示された結果と同様の傾向がみられ，予備実験に示すように推定値の分散が小さい場合によい推定精度を示すことが確認できた．

ここで，実際のゲームに適用する際には完全調査のサンプル数は指定できず，閾値とそのときの状況に応じて完全調査の列挙数が変化する．図 12 の結果より，完全調査のサンプル数が 10〜60 程度であれば有効な精度の向上が望めることが分かる．完全調査のサンプル数が少ない方が破棄するサンプルが少なくなり無駄が減少することを考慮して，完全調査のサンプル数が 10〜20 程度になれば効率が良いことが分かる．そこで，200,000 サンプルに関して調査した結果，閾値を 0.01 と指定した場合に完全調査に含まれることになるサンプル数は平均約 14 個であることが分かり，この閾値設定が妥当であったことが確認できた．

5.3 実　験　2

次に，Dummy-Agent7 体を対戦相手に MC-Agent，PEXI-MC-Agent，MAX-Agent のそれぞれが 100 ゲーム行った場合のスコアの違いを比較する．**図 13** はそれぞれ 100 ゲームにおけるスコアの平均値，標準偏差，最大値，および最小値を示したグラフであり，縦軸がスコアを表す．期待値を使用していない

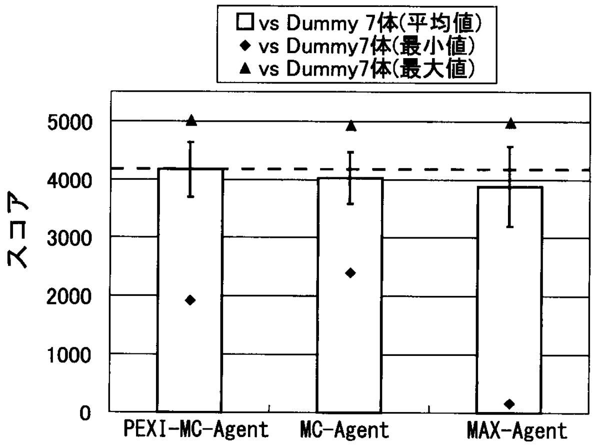


図 13 100 ゲームの結果
Fig. 13 Results in 100 games.

MAX-Agent と比較して，MC-Agent および PEXI-MC-Agent の方が平均スコアが上昇し，標準偏差も小さくなっていることが分かる．同時に，それぞれ最小スコアも改善されており，最大スコアも低下していない．ゲームの状況を考えると，期待値を導入することによって高スコアの無理な旅行パック実現に失敗してスコアを落とすことを回避できたためだと考えられる．また，MC-Agent よりも PEXI-MC-Agent の方がスコアが良い．期待値の推定精度の向上がスコアの向上に寄与するかは対象問題によるが，TAC CLASSIC では PEXI-MC 法の導入が有効であったことが確認できた．

また今回の実験では，入札とその結果という単純な確率事象のとらえ方であり，さらに期待値を推定するために用いた生起確率は事前実験によるもので正しい値ではないにもかかわらず，期待値の導入によってスコアの改善が行えることが確認できた．つまり，実際のゲームの状況が意思決定に応じて完全に確率的に生起する場合でなくても，期待値の導入によって不確実性を考慮することでより良い意思決定が行える可能性を示すことができた．

6. 終わりに

本論文では，高い生起確率を持つ結果を列挙することが比較的容易である場合に MC 法による期待値の推定精度を向上させる PEXI-MC 法を提案した．まず，提案手法による精度向上の可能性を式展開により示し，予備実験によってその有効性を示した．予備実験の結果から，生起確率の偏りが大きい場合に，提案手法が効果を発揮することを確認した．さらに，応用の一例として TAC CLASSIC におけるエージェントの意思決定へ提案手法を適用した．実験結果から，提案手法を導入することで期待値推定の精度が向上し，ゲームのスコア向上がはかれることが確認された．

今回の提案では列挙のための閾値は予備実験により決定したが，提案法の精度や計算コストは閾値に影響されるため，列挙を行いながら自動調整することが望ましい．今後は，閾値の自動調整について提案を行う予定である．

謝辞 本研究を進めるにあたり多大なご協力，ご助言をくださった，産業技術総合研究所情報技術研究部門の車谷浩一氏に深く感謝いたします．

参考文献

1) 小林康幸，坂口敏洋，傍島教文，難波伸也，黒岩大史：期待値を用いたゲーム木探索，情報処理学会論文誌，Vol.43, No.10 (2002).
2) 竹蔵正史，橋本 強，梶原羊一郎，長嶋 淳，飯田弘之：コンピュータ将棋における実現確率探索の研究，ゲームプログラミングワークショップ 2002，ISBN1344-0640, pp.87–92 (2002).
3) 福島孝治：モンテカルロ法の前線—サイコロ振って積分する方法，若手研究者・学生向けに最新技術をわかりやすく紹介する講演会「確率的アルゴリズムによる情報処理」講義資料，京都 (2003).
4) 津田孝夫：モンテカルロ法とシミュレーション，培風館，東京 (1977).
5) Wellman, M.P. and Wurman, P.R.: A Trading Agent Competition for the Research Community, *IJCAI-99 Workshop on Agent-Mediated Electronic commerce*, Stockholm (1999).
6) Wellman, M.P., Greenwald, A., Stone, P. and Wurman, P.R.: The 2001 Trading Agent Competion, *Revised version of 14th Conference on Innovative Applications of Artificial Intelligence*, Edmonton, pp.935–941 (2002).
7) Greenwald, A.: The 2002 Trading Agent Competition: An Overview of Agent Designs, *AI magazine* (Apr. 2003).
8) Trading Agent Competition homepage (2001–2003). http://www.sics.se/tac/

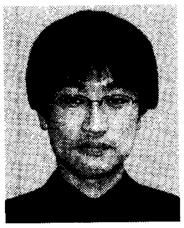
(平成 16 年 10 月 6 日受付)
(平成 17 年 10 月 11 日採録)

推 薦 文

部分的完全調査モンテカルロ法の提案と，不確実性を考慮に入れたエージェントの意志決定設計への応用が可能であることを実験により示した点が評価できる．
(情報処理北海道シンポジウム 2004 プログラム委員
鈴木恵二)

川村　秀憲（正会員）

1973 年 5 月 3 日生．1996 年北海道大学工学部情報工学科卒業，2000 年同大学大学院工学研究科博士後期課程期間短縮修了．同年同大学院工学研究科助手．2004 年同大学院情報科学研究科助手，現在に至る．飛行船ロボット，マルチエージェントシステム，複雑系工学，観光情報学等の研究に従事．博士（工学）．人工知能学会，日本オペレーションズ・リサーチ学会，観光情報学会，電気学会各会員．



小野寺将輝

1981 年 1 月 2 日生．2003 年北海道大学工学部情報工学科卒業，2005 年同大学大学院工学研究科修士課程修了．同年株式会社 NTT ドコモ北海道入社．現在に至る．マルチエージェントシステム等の研究に従事．

大内　　東（正会員）

1945 年 8 月 19 日生．1968 年北海道大学工学部応用物理学科卒業，1974 年同大学大学院工学研究科博士後期課程修了．同年同大学工学部助手，助教授，1989 年同大学大学院工学研究科教授．2004 年同大学院情報科学研究科教授，現在に至る．飛行船ロボット，DNA コンピューティング，マルチエージェントシステム，医療システム等の研究に従事．工学博士．日本オペレーションズ・リサーチ学会，人工知能学会，医療情報学会，観光情報学会各会員．