社団法人　情報処理学会　研究報告
IPSJ SIG Technical Report

2004－ＩＣＳ－135　（３）
2004／3／15

部分的全調査モンテカルロ法による不確実性を考慮した行動評価関数の設計とTACへの適用

小野寺将輝†　川村　秀憲†　車谷　浩一††　大内　東†

† 北海道大学大学院工学研究科　〒060-8628 北海道札幌市北区北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: †{onodera,kawamura,ohuchi}@complex.eng.hokudai.ac.jp, ††k.kurumatani@aist.go.jp

あらまし　電子商取引などの分野におけるソフトウェアによる自動的な意思決定に関して，期待値を用いた行動評価関数を設計する際の，期待値の推定の精度を向上させる一手法として部分的全探索モンテカルロ法を提案し，実験によってその有効性を検証した．また一例として，設計した評価関数を TAC CLASSIC へ適用して実験を行い，その有効性が確認された．
キーワード　モンテカルロ法，Trading Agent Competition，行動評価関数

Design of action-value function based on Monte Carlo method with partial exhaustive investigation, and Application to TAC

Masaki ONODERA†, Hidenori KAWAMURA†, Koichi KURUMATANI††, and Azuma OHUCHI†

† Graduate School of Engneering, 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: †{onodera,kawamura,ohuchi}@complex.eng.hokudai.ac.jp, ††k.kurumatani@aist.go.jp

Abstract　About automatic decision-making by the software in fields, such as electronic commerce, the action-value function, which is in consideration of the uncertainty using the expected value, was designed. Moreover, as a way method which raises the accuracy of calculation of an expected value, Monte Carlo method with partial exhaustive investigation was proposed, and the validity was verified by experiment. At the end, the action-value function was applied to TAC CLASSIC, and the validity was checked.
Key words　Monte Carlo method, Trading Agent Competition, action-value function

1. はじめに

近年，電子商取引やコンピュータゲームなどの分野において，ソフトウェアによる自動的な意思決定に関する研究が行われている．例えば，Trading Agent Competition という国際フォーラムにおける自動取引に関する研究 [1]～[3] や将棋のコンピュータプログラムの研究 [7] などがある．これらのような意思決定の研究に共通する課題として，意思決定を行う際に各行動をどのように評価するかという点が挙げられる．意思決定の際には，選択可能な行動が多数存在し，さらにそれぞれが複数の異なる結果を導く可能性を持つ場合が多い．つまり，ある行動を選択したときに，そこから導かれる結果を正確に予測することができないという不確実性を含んでいる場合があると考えられる．

ここで，行動選択に対して確率的に結果が生起し，さらにその確率が既知である場合を仮定する．このとき結果への不確実性を考慮する一つの方法として，生起する可能性のある全ての結果の評価値から計算される期待値を行動評価関数として用いることが考えられる．しかし，考慮しなければならない結果が膨大になると，計算コストに対する制約などによって，実際に全ての結果についての期待値を計算することが不可能となる場合があると考えられる．そのような場合，何らかの方法を用いて期待値を推定する必要が出てくる．期待値の推定には，モンテカルロ法が用いられることが多い [6]．モンテカルロ法による推定の精度は高いことが望ましいが，計算コストの制約により多数のサンプルを調査することができない場合がある．そのため，少数のサンプルで可能な限り精度を上げたいという，従来のモンテカルロ法に対する要求とは多少異なる要求が生じる．

さてここで，対象とする問題によっては，生起確率の高い結

```
calc(int target, int cand){
  for(int a = 0; a < j_cand; a++){
    double temp = target · p(x^a_cand);
      if(temp ≥ threshold){
        if(cand == n){
          事象の組合せを保存;
        }
        else{
          calc(temp, cand + 1);
} } } }
```

図 1　列挙のアルゴリズム

果を列挙することが，比較的簡単にできるという特徴を持つ場合がある．このような場合には，この特徴を利用することで，期待値に大きく影響を与えている部分を効率よく抽出し，期待値の推定の精度を向上できる可能性があると考えられる．

そこで本研究では，このような特徴を利用し，モンテカルロ法を用いた期待値の推定の精度を向上させる手法を提案する．以下では，まず対象とする問題を説明したのち，手法の提案を行う．さらに，提案手法を実際の問題に適用して評価を行う．

2. 部分的全調査モンテカルロ法の提案

2.1 対象とする問題の設定

ある一つの行動を選択することによって一つの結果が確率的に生起するとする．ただし，生起する可能性のある結果は複数存在するような場合を考える．このときその行動の評価値として，実現の可能性のある全ての結果の評価値から計算される期待値を用いることを考える．つまり，生起する可能性のある結果の集合を X とし，ある結果 $x(\in X)$ の生起確率を $p(x)$，評価値を $f(x)$ とすると，その行動の評価値は，

$$E = \sum_{x \in X}(f(x) \times p(x)) \tag{1}$$

として計算出来る．

ここで，N 個の複合事象 X_i $(i = 0, 1, \cdots, N)$ が存在し，それらの積事象として X が定義される場合を考える．つまり $X = X_0 \cap X_1 \cap \cdots \cap X_N$ である．また，各事象は $X_i = \{x_i^0, x_i^1, \cdots, x_i^{n_i}\}$ (n_i：事象 i ごとの任意の定数) という複数の要素から成るとする．つまり，ある結果 x は，$x = x_0^1 \cap x_1^1 \cap \cdots$ のように定義されることになり，要素の組合せが少しでも異なれば，異なる結果が導かれることを意味する．またこのことより，事象 X_i の各要素の生起確率 $p_i(x_i^{j_i})$ (j_i : n_i 以下の任意の正の整数) が得られると仮定すると，ある結果 x の生起確率 $p(x)$ は，$p(x) = p_1(x_0^1) \times p_2(x_1^1) \times \cdots$ という各要素の生起確率の積として計算することができる．

以上のような特徴を持った問題を対象として考えることとし，以下の議論を行っていく．

2.2 生起確率が高い結果の列挙の方法

前述のように複数の事象 X_i の積事象として各結果の生起確率が計算できるような場合，全体を調べなくても，ある閾値を決めてそれより生起確率の高い結果を列挙することができる．図 1 にそのアルゴリズムを示す．

計算は，$target = 1.0, cand = 0$ を初期値として $calc$ を実行することで行う．$calc$ は，積事象の生起確率が閾値以上のもののみを結果として保存し，計算過程で閾値未満となる要素の組合せに関しては，その先の計算を省略するという枝刈りを行う．従って，閾値をある程度高めに設定すれば，省略できる積事象の数が増え，列挙に要する時間を短縮することが可能となる．

2.3 部分的全調査モンテカルロ法

ここで，ある程度生起確率が高い結果を列挙できるという特徴を利用し，その部分の全調査を組み合わせることでモンテカルロ法の精度を上げる手法を提案する．以下，この手法を部分的全調査モンテカルロ法と呼ぶこととする．ここでモンテカルロ法の精度は，得られる期待値の分散によって評価できるので，提案手法の評価は，加重サンプリング初等的モンテカルロ法を用いた場合の期待値の分散との比較によって行うことができる．

まず加重サンプリング初等的モンテカルロ法を式で表す [5]．発生させる総サンプル数を M とし，$p(x)$ に従って発生させた M 個の乱数を $\xi_1, \xi_2 \cdots, \xi_M$ とすると，期待値 E_1 とその分散 $var E_1$ は，

$$E_1 = \frac{1}{M}\sum_{i=1}^{M} f(\xi_i) \tag{2}$$

$$var E_1 = \frac{1}{M} var f(\xi) \tag{3}$$

となる．次に部分的全調査モンテカルロ法を式で表す．まず結果の集合 X を，生起確率が閾値より大きい結果の集合 $X^{(1)}$ と，それ以外の集合 $X^{(2)}$ に分割する．このとき $X^{(1)}$ に含まれる要素数を $|X^{(1)}|$ とする．ここで，それぞれの集合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 $M_1 = |X^{(1)}|$ である．$X^{(1)}$ に関する全調査結果を E_1' とすると，

$$E_1' = \sum_{x \in X^{(1)}} f(x)p(x) \tag{4}$$

となる．$X^{(2)}$ に関しては加重サンプリング初等的モンテカルロ法で計算する．$x(\in X^{(2)})$ の生起確率の分布を $p_2(x)$ とすると，

$$p_2(x) = \frac{p(x)}{\sum_{x \in X^{(2)}} p(x)} \qquad (x \in X^{(2)}) \tag{5}$$

となる．$p_2(x)$ に従う $M_2(= M - M_1)$ 個の乱数を $\xi_1^{(2)}, \xi_2^{(2)}, \cdots, \xi_{M_2}^{(2)}$ とすると，モンテカルロ法による出力 E_2' は，

$$E_2' = \frac{1}{M_2}\sum_{i=1}^{M_2} f(\xi_i^{(2)}) \tag{6}$$

となる．このときこの手法によって得られる期待値 E_2 とその分散 $var E_2$ は，

$$E_2 = E_1' + \sum_{x \in X^{(2)}} p(x) E_2' \tag{7}$$

$$var E_2 = \frac{(\sum_{x \in X^{(2)}} p(x))^2}{M_2} var f(\xi^{(2)}) \tag{8}$$

となる．これらから，

$$var E_1 > var E_2 \tag{9}$$

を満たす場合に精度が向上することがわかる．次に実験によって精度の違いを考察する．

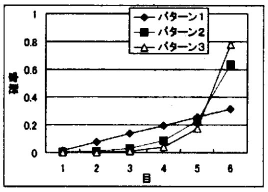


図 2 各日の生起確率（3パターン）

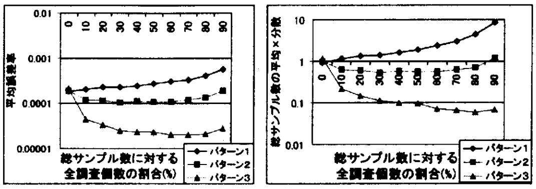
2.4 実験設定

部分的全調査モンテカルロ法の導入による精度の変化を調べるために，シミュレーションによる加重サンプリング初等的モンテカルロ法との比較を行う．実験設定として，複合事象の最も単純な例の一つであるサイコロを複数ふることを考える．このときサイコロの日の組合せごとにランダムな評価値を割り当て，その評価値の期待値を求めることとする．また生起確率の分布の偏りによる結果の違いを比較するために，図 2 のような，各日の生起確率が異なる 3 パターンを用意する．ある程度組合せ数が多くなるようにサイコロの数は 8 個とし，このとき総パターン数は 1,679,616 通りとなる．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 は総パターンの約 0.1%である 1,680 個とする ($M = 1680$)．またこのとき，2 つに分けた集合のうち，モンテカルロ法で推定する部分に含まれるサンプルのみを発生させることが現時点では困難であるため，加重サンプリングと同様の方法でサンプルを発生させ，モンテカルロ法で推定する部分に含まれるサンプルのみを使用することとし，それ以外は使用しないことにする．最終的に全調査を行うサンプル数 M_1 との合計が 1680 個になるまでサンプルを発生させる．評価値は，平均 100，分散 1 の正規分布に従った乱数によって割り当てる．実験 1 としてある 1 種類の評価値の割り当てに関して，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 1000 試行の推定を行って誤差を比較する．実験 2 として評価値の割り当てを 10 種類用意しそれぞれ 100 試行ずつ，合計 1000 試行行って誤差の比較を行う．またこのとき全調査を行う個数 M_1 をより細かく総サンプル数 M の 5%刻みで変化させる．

2.5 結果・考察

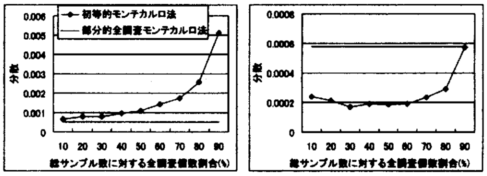
図 3(a) は，実験 1 における 1000 試行の平均誤差率である．各日の生起確率の偏りが小さいパターン 1 では全調査を加えることで誤差率が増加してしまっている．しかし，パターン 2，3 と偏りが大きくなるにつれて，全調査を加えることで明らかに初等的モンテカルロ法よりも精度が高くなる部分があることが確認できる．ただし，今回の実験では部分的全調査モンテカルロ法を行うにあたって，発生させたのに使用しない無駄なサンプルが発生している．そこで，使用しなかったサンプルに関するコストも考慮するために，(結果の分散 × 発生させた総サンプル数) という値を図 3(b) に示す．グラフから，コストを考慮しても，初等的モンテカルロ法より高い精度が得られる場合があることがわかる．

図 4 は，実験 1 における各パターンごとの 1000 試行の分散

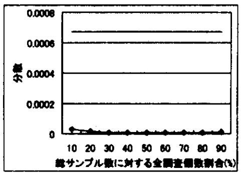


(a) 平均誤差率　(b) 総サンプル数の平均×分散

図 3 実験 1 の結果



(a) パターン 1　(b) パターン 2



(c) パターン 3

図 4 実験 1 における分散の変化

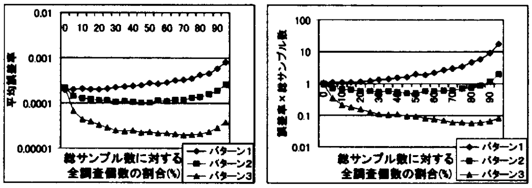
である．直線が基準となる初等的モンテカルロ法での分散を表し，折れ線は全調査に使用するサンプル数を変化させた時の分散の変化を表している．パターン 1 では初等的モンテカルロ法の方が，2 と 3 では部分的全調査モンテカルロ法の方が分散が小さくなっており，これより式 (9) を満たす場合に精度が向上していることがわかる．

図 5 は，実験 2 における結果のグラフである．評価値の割り当て方を何パターンか変化させても，実験 1 と同様の傾向が見られる．つまり，生起確率に偏りが存在し，全調査される部分の生起確率が全体に占める割合がある程度高い場合には，部分的全調査モンテカルロ法によって精度を高めることができると考えられる．これらの結果を踏まえて，次に Trading Agent Competition への適用を行う．

3. Trading Agent Competiton (TAC)

3.1 TAC とは

Trading Agent Competition (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の研究を行っている国際フォーラムであり，独自のシミュレーションシナリオを作成して研究を行っている．現在は，電子市場での連続オークションにおいて複数の代替財・補完財の組合せを自動的に取引するエージェントの研究を行っている TAC CLASSIC と，サプライチェーン・マネジメントにおけるエージェントによるプランニングに関する研究を行っている TAC SCM という 2 種類の



(a) 平均誤差率　　(b) 10 パターンの (総サンプル数の平均×分散) の平均

図 5　実験 2 の結果

シナリオが用意されており，それぞれのシナリオへ複数の研究者が自分の設計したエージェントを持ち寄って競い合うことで，より高品質な研究を行うことを目指している．

本研究では，このうちの TAC CLASSIC を対象として，エージェントの意思決定へ先ほどの評価関数を導入する．このシナリオにおいて，エージェントは複数の商品に対して入札を行うが，必ずしも入札した商品を全て入手できるわけではない．また事前に正確な価格を知ることは不可能であり，他人がどのような行動をとっているかということを正確に知ることも不可能である．これらのような困難な問題を多数含んでおり，このシナリオは不確実性を含んでいると考えることができる．以下では，TAC CLASSIC の概要と設定の説明を行う．

3.2　TAC CLASSIC

このシナリオにおいて各エージェント（参加者）は，それぞれに個別に割り当てられた８人の客の好みに合わせた５日以内の旅行パックを計画し，必要な複数の商品を入手する．エージェントの目的は，客の満足度を出来る限り最大化することである．以下，このシナリオのシミュレーション１試行を１ゲームと呼ぶ．１ゲームには８体のエージェントが参加し，１２分で行われる．旅行パックは，往復のフライトチケット，その間のホテルの予約，娯楽チケットの３種類の商品で構成され，それぞれ異なる市場で取引される．以下でそれぞれの説明を行う．

フライトチケット

single seller auction で取引されエージェントは買うのみである．１日目から５日目までそれぞれ行きと帰りのチケットが個別に存在する．ただし，客は少なくとも一日は滞在しなければならないので，１日目の帰りと５日目の行きのチケットは存在しない．また全てのチケットに関して売り切れはない．価格は，内部的に事前に決められた関数に従って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がある．

ホ テ ル

２種類のホテル（TT，SS）が存在し，TT は SS よりも良いホテルという設定である．旅行中に異なるホテル間を移動することはできない．各ホテルごとに一日１６部屋が用意されている．ただし，最終日は帰るのみなのでホテルは存在しない．それぞれのホテル・日にちごとに異なる ascending multi-unit auction で取引されるので，計８つの市場が存在することになる．これらの市場は，４分から１１分の間に毎分一つずつランダムに終了され，その瞬間に清算される．落札価格は１６番目に高い入札額となり，一度行った入札は自分で取り消すことはできず，一度入手した商品は売ることはできない．

娯楽チケット

３種類のチケット（ワニレスリング，遊園地，博物館）が，１日目から４日目までそれぞれ各日８枚ずつ存在し，滞在中の出発日以外に一日一枚だけ使用できる．ゲーム開始時に各エージェントに一定枚数がランダムに配布され，continuous double auction においてエージェント間で取引を行う．価格は需要と供給のバランスで決定される．

各客は，以下のような個別の好みを持っている．

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

ここから客ごとの効用（U）が以下のように計算される．

$$U = 1000 - travel\_penalty + hotel\_bonus + ent\_bonus$$
$$travel\_penalty = (|AA - PA| + |AD - PD|)$$
$$hotel\_bonus = TT? \times HP$$
$$ent\_bonus = AW? \times AW + AP? \times AP + MU? \times MU$$

ここで $TT?$ はその客が TT へ泊まったかどうか，$AW?, AP?$, $MU?$ は各娯楽チケットが含まれているかどうかを（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コストの総和を引いたものとして計算される．

このシナリオを使用したシミュレーション実験を行うにあたって，実験に使用する３種類のエージェントを用意する．以下でそれぞれに関して説明を行う．

4.　ExpAgent の設計

ExpAgent は，TAC CLASSIC をスコアの最大化問題と捉え，それを満たすためにゲーム中の各段階において，価格の予測，注文の決定，入札という３つの動作を行う．また，注文の決定は GA によって行うが，その際の適応度の計算に先ほど設計した評価関数を導入する．以下でそれぞれの動作の詳細，及び適応度の計算方法を述べる．

4.1　価 格 予 測

- フライトチケット：価格予測はせず現在価格を使用する．
- ホテル：ニューラルネットワーク (NN) を用いて１分後の価格を予測する．NN は３層 (入力層，隠れ層，出力層) から構成され，各層はそれぞれ３つ，８つ，１つのユニットを持つ．2 分前，1 分前，現在の価格を入力とし，1 分後の価格を出力する．各重みは，TAC2003 の Qualify Round における 590 ゲームでの価格変化のデータを使用して，真値と出力値の二乗誤差が最小になるように GA で事前に固定しておいたものを使用する．また，各市場ごとに個別の NN を用意する．
- 娯楽チケット：価格予測はせず現在価格を使用する．

4.2　注文決定方法

注文は GA で計算する．GA の設計は TAC 参加チームの１つ Painin NEC を参考にし [3]，いくつかの変更を加えた．１人

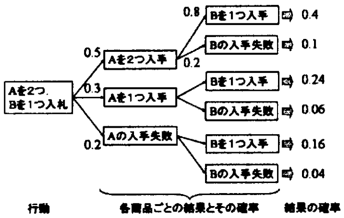


図 6　行動と結果の分岐例

の客の旅行パックを長さ 5 の文字列 $te_1e_2e_3e_4$ で表現する．t は -1〜19 の値をとり，それぞれが異なる日程とホテルの種類を表す．例えば，$t = 0 or 10$ の時は 1 日目到着で 2 日目出発，$t = 1 or 11$ の時は 2 日目到着で 3 日目出発，$\cdots$，といった具合である．$t = -1$ は，その客に関しては旅行パックを作成しない，つまり何も商品を購入しないことを示す．また，$0 < t < 9$ の時は SS へ，$10 < t < 19$ の時は TT へ泊まることを示す．$e_1e_2e_3e_4$ は 0〜4 の値をとる．その値が旅行パックに含まれる娯楽チケットの種類を表し，添え字が日付を表す．例えば $e_1 = 1$ は 1 日目のワニレスリング，$e_1 = 2$ は 1 日目の遊園地，$e_1 = 3$ は 1 日目の博物館を表し，$e_1 = 0$ は 1 日目の娯楽チケットは含まれないことを表す．1 個体は，8 人分の長さ 40 の文字列に，後に説明する適応度計算時に使用する 8 人の客の優先順位を示す長さ 8 の文字列を加えた，長さ 48 の文字列によって表現される．遺伝的操作として，エリート保存戦略を併用したトーナメント戦略，1 点交差，突然変異を行う．各種パラメータは，集団数 200，世代数 100+1，交叉確率 0.9，突然変異確率 0.05 とする．GA による計算は，ゲーム開始時と 3〜11 分の間の毎分 1 回，計 10 回行われる．この GA によって決定された旅行パックに従い，そのために必要な商品に対して入札を行う．

4.3 GA における適応度計算

ExpAgent は，GA の適応度として，その個体が示す旅行パック実現に必要な商品へ入札を行った際の，得られるスコアの期待値を用いる．従来研究では，必要な商品は全て手に入ると仮定して計算したスコアを適応度としていた．しかし本研究では，いくつかの商品を入手できなかった場合も考慮することで，不確実性を考慮した意思決定の実現を目指す．

例として図 6 に示すような，商品 A を 2 つと商品 B を 1 つに対して入札するという行動を考える．このときそれぞれの商品ごとに独立に，いくつ入手できるかという確率が決まっているとすると，可能性のある 6 種類の結果についての生起確率がそれらの積事象として得られる．この確率とそれぞれの結果のスコアから計算した期待値がその行動の評価値，つまり，そのような行動を表す個体の適応度となる．

具体的には，まず予備実験として後述する DummyAgent を相手に 800 ゲームを行う．この際，注文をランダムに決定し，各商品ごとに注文数と実際の落札数のデータを得る．ここから，全ての商品の注文数と落札数の組に関して，「a 個注文したときに b 個得られる確率」を計算する．この確率と，入手できる商品数の組合せごとに計算されるスコアから期待値を計算し適応度として使用する．また，いくつか必要な商品を入手できない状況を仮定したときのスコアは，個体が示す優先順位に従って商品を割り当て，その時のスコアとして計算される．

ここで，考えられる全ての状況を計算し期待値の真値を求めることは，リアルタイム性が求められるゲームの性質から不可能であると考えられる．そこで 2.3 節で説明した部分的全調査モンテカルロ法を用いて，期待値を推定することとする．このとき 2.5 節で行った実験とは異なり，全調査を行う部分を個数で指定することが困難である．そこで閾値を 0.01 と設定し，それより生起確率が高い部分は全調査を，それ以外の部分はモンテカルロ法による推定を行うこととする．図 6 のように，各商品ごとの確率の積事象として各結果の生起確率が得られるので，生起確率の計算過程で閾値を下回った場合にはそれより先の計算を行う必要がなくなるため，全調査にはそれほど多くの計算時間はかからない．また推定に用いるサンプルとして，各商品ごとの確率に従って，0〜注文数 までの乱数を発生させて入手できた商品数を仮定し，その時の生起確率とスコアを計算する．本研究では，過去の実験におけるモンテカルロ法に費やせる計算時間の経験から，発生させる総サンプル数は 100 とする．つまりモンテカルロ法に用いるサンプル数は，(100 − 全調査の数) となる．また，ランダムに発生させたサンプルが全調査部分に含まれる場合には，代わりのサンプルを発生させることは行わず，そのサンプルは使用しない．つまり結果的に使用するサンプルが 100 個を下回る場合もあることになる．

4.4 入札規則

注文を決定後，以下の規則に従って入札を行う．

- フライトチケット：必要な商品は全て即購入する．
- ホテル：1 分ごとに，予測価格＋上乗せ価格 α で入札する．この α は，予測価格があくまで 16 番目に高い入札額であるため，より確実に必要な数を手に入れるためのゆとりを持たせることを意味する．
- 娯楽チケット：事前に決めた価格を決定する関数に従って，買い注文は時間の経過に従って少しずつ高く，売り注文は少しずつ安く更新する．

5. MaxAgent の設計

基本的には全て ExpAgent と同様の行動を行うが，GA における適応度計算のみが異なる．MaxAgent は，ある個体が表す旅行パックに必要な商品は全て手に入ると仮定し，その時のスコアを適応度として使用する．つまり不確実性を考慮していない意思決定と考えられる．

6. DummyAgent

[4] において配布されている最も基本的なエージェントであり，このプログラムに変更を加えることで前述の 2 種類のエージェントを実装している．入札額などいくつかのパラメータに変更を加えて使用する．具体的には以下の行動をとる．

- 注文の決定：価格は考慮せず，客の好み通りの日程とする．ホテルは HB が 70 以上なら TT，それ以外は SS を使用する．娯楽チケットはボーナスの高い順に日程に含めるだけ含む．
- 入札規則：フライトチケットは全て即購入する．ホテル

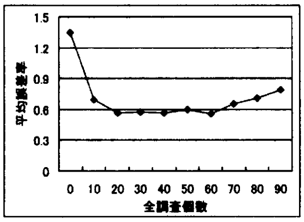


図 7　200000 サンプルに関する平均誤差率

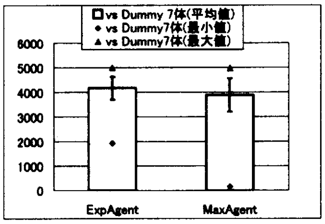


図 8　100 ゲームの結果

は最初は 200 ドルで入札し，以降は 50 ドルずつ競り上げる．娯楽チケットは，他の 2 種類のエージェントと同様に，事前に決めた関数によって決定される価格で入札する．

7. 実　　験

7.1 実験 1 設定

まず，ExpAgent に実装した部分的全調査モンテカルロ法の効果を調べるための実験を行う．実際に 1 ゲームを行い，そのときの GA で発生した 20000 個 × 10 回 = 200000 個 の個体に関して，加重サンプリング初等的モンテカルロ法，全調査を行うサンプル数を 0〜90 まで 10 個 刻みで変化させた部分的全調査モンテカルロ法の 2 種類の手法によって期待値を推定し，全調査を行った結果との比較を行う．

7.2 結果・考察

図 7 は 200000 個の個体に関して，部分的全調査モンテカルロ法によってスコアの期待値を計算するときに，全調査個数を変化させた場合の平均誤差率の変化のグラフである．全調査個数が 0 のときが通常の加重サンプリング初等的モンテカルロ法の結果にあたる．明らかに全調査を加えることで誤差率を減少することができていることが見て取れる．また実際のゲームを行う場合には，全調査の個数ではなく閾値を指定することになるが，今回使用した 200000 サンプルに関して閾値を 0.01 に設定したときに全調査部分に含まれる個数の平均を計算したところ，約 14 個であった．この結果から，設定した閾値がある程度適切であり，誤差率を減少できることが確認された．

7.3 実験 2 設定

次に ExpAgent，MaxAgent のそれぞれを DummyAgent7 体と 100 ゲームずつ対戦させ，行動の評価関数の違いによるゲームの結果の変化を比較する．また，確率に従って行動しているわけではない DummyAgent とのゲーム結果を確率的に捉え，それを意思決定に組み込むことの有効性も検証する．

7.4 結果・考察

図 8 は，それぞれのエージェントの 100 ゲームにおける平均スコア，標準偏差，最大スコア，最小スコアを示したグラフである．明らかに平均スコアが向上し，標準偏差が抑制されていることがわかる．最小スコアを比べてみると，ExpAgent の方は低いスコアをとることを回避できていると考えられる．また，最大スコアが下がってしまうといった現象は起きていない．これらの結果から，行動の評価関数として期待値を使用することで，無理な旅行パックの実現を目指してスコアを落とすことを回避できていると考えられ，不確実性を考慮した意思決定がある程度実現されていると考えることができる．また，DummyAgent とのゲーム結果から計算した確率を導入することで，スコアの向上と標準偏差の抑制が見られたことから，本実験の設定上はエージェントの行動の結果を確率的に捉えることが妥当であると考えられる．

8. 結　　論

不確実性を考慮した評価関数として期待値を用いる際に，真値を計算することが困難である場合が多い．そこで本研究では，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という特徴を持つ問題が存在することに着目し，その特徴を利用した期待値の推定の精度を上げる手法として部分的全調査モンテカルロ法を提案した．その有効性を式による証明とシミュレーションによる実験で確認した．さらに，応用例として TAC CLASSIC にその評価関数を導入した実験結果から評価関数の有効性が確認された．

TAC CLASSIC への応用においては，今回は予備実験によってある程度正確な生起確率が得られたが，適用対象が変わった場合に，どのようにして正確な生起確率を見積もるかが今後の課題として挙げられる．また，TAC CLASSIC 以外にも適用し，その有効性を検証する必要がある．

文　　献

[1] Michael P. Wellman, Peter R. Wurman, A Trading Agent Competition for the Research Community, IJCAI-99 Workshop on Agent-Mediated Electronic commerce, Stockholm, 1999.

[2] Michael P. Wellman, Amy Greenwald, Peter Stone, Peter R. Wurman, The 2001 Trading Agent Competion, Revised version of Fourteenth Conference on Innovative Applications of Artificial Intelligence, pages 935-941, Edmonton, 2002.

[3] Amy Greenwald, The 2002 Trading Agent Competition:An Overview of Agent Designs, AI magazine, April, 2003.

[4] Trading Agent Competition WWW homepage, http://www.sics.se/tac/, 2001–2003.

[5] 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.

[6] 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, Oct.2002.

[7] 将棋プログラム「激指」のページ，http://www.logos.t.u-tokyo.ac.jp/ gekisashi/ ,