

# ターン制戦略ゲームにおけるファジィ評価を用いた探索木の枝刈り

Cutoff Using Fuzzy Evaluation on AI of Turn-based Strategy Games

武藤 孝輔                      西野 順二  
Kohsuke Mutoh                Junji Nishino  
電気通信大学

The University of Electro-Communications

概要: ターン制戦略ゲームのような探索空間の巨大なゲームは囲碁や将棋にくらべ十分な研究がなされておらず, 人工プレイヤーの AI は人間プレイヤーよりも弱い。ユニットを動かす順序とそれぞれの行動の組み合わせで, 同一ターン内の合法手の数が膨大であるが, すべてのユニットの行動の組み合わせ全体を探索するのではなく, 1つのユニット行動ごとに分解しゲーム木を構成してノードごとの分岐を減らすことが可能である。ノードごとの分岐を減らしても合法手の数は膨大なままであるため, ゲーム木から悪手を枝刈りするか, 有望そうなノードを優先的に探索することで探索の更なる効率化が期待できる。そこで, 本研究では人間的直感に基づいた判断による悪手の枝刈りおよびノードの探索優先度の順位付けを行うためにファジィ評価を用い, その有効性を検討した。

## 1 はじめに

ターン制戦略ゲーム [1] は, 戦争を模したボードゲームで, 同一ターン中に複数の駒を動かす複数着手性のため探索空間が広く人工プレイヤーを作るのが困難な対象である。また囲碁や将棋にくらべ, ターン制戦略ゲームの人工プレイヤーの AI は人間プレイヤーよりも弱く, 上級者の対戦相手が務まらない。

本研究は, ターン制戦略ゲームを具体的な対象として, 1ターンの合法手の多いゲームの人工プレイヤーアルゴリズムに対し, 人の直感的な判断をファジィ評価によって取り入れることを目的とする。

ターン制戦略ゲームの1ターンは, 複数のユニット(駒)を動かすことで成り立っている。このため, ユニートを動かす順序とそれぞれの行動の組み合わせで1ターン全体のプレイヤーの手の可能性が  $10^3$  以上になることが多く膨大となっている。この構造をもとに, ターン単位での合法手, すなわちすべてのユニットの行動の組み合わせ全体を探索するのではなく, 1つのユニット行動ごとに分解しゲーム木を構成して探索することを考える。本研究ではこうして再構築したゲーム木をユニット行動木と呼び, このユニット行動木の上でモンテカルロ木探索 (UCT) による探索を行なう。

こうした, 探索の効率化はターン制戦略ゲームにおいての重要な課題のひとつであり, これまでに, シミュレーションを途中で打ち切りその局面を状態評価関数で評価するモンテカルロ法 [2] や, 選択枝の枝刈りを行う UCT 探索を用いた手法 [3] が提案されている。

ユニット行動木 UCT 探索を行なっても, なお探索木の分枝数は, その他のゲームと比べても多く, 探索の深

さは実際には1ターン分にも満たないことが多い。例えばユニット数が6ずつあるとき, 通常の意味で12手先まで探索してはじめてやっと相手の対応までが反映される。人間はこうした膨大な局面展開を直感的に評価して有効な着手を選択することができ, 人工プレイヤーよりも強い。この人間の直感評価を多次元空間でのファジィ集合 [4, 5] によりモデリングし, 探索木の前向き枝刈りおよび UCT 探索の優先度に使用することを検討する。

## 2 TUBSTAP

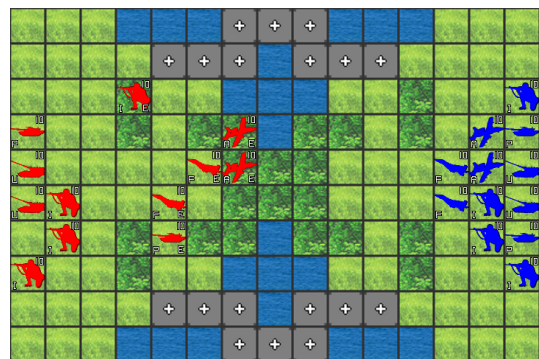


図 1: TUBSTAP のゲーム画面

TUBSTAP[1] は北陸先端科学技術大学院大学 (JAIST) 池田研究室がターン制戦略ゲームの思考アルゴリズムの研究・開発の為に立ち上げたプロジェクト, およびその対象ゲームのことである [6]。ゲームのルールは簡略化されており, 多くの戦略ゲームに共通する基本的な要素だけを持った設定となっている。本研究ではこのゲームを具体的な対象として実験を行なった。

将棋やチェスの場合、駒  $u$  が  $j$  個、駒  $u$  の移動先が  $m_u$  個あった時、その手番内で取りうる手の数は、 $\sum_{i=1}^j m_{u_i}$  である。

一方、複数着手性のある TUBSTAP の場合、この手の数は  $\prod_{i=1}^j (m_{u_i} + 1) \times j!$  程度となる。移動順が異なることで同一の局面となることも、ならないこともある。

将棋やチェスが各駒の移動先の総和であるのに対し、TUBSTAP は各駒の移動先と移動しないという選択肢の総乗と駒を動かす順番の総数との積にまで手の数が増えている。このため、全てを探索することは現実的には不可能である。

## 2.1 ゲームのルール

TUBSTAP は相手の全てのユニットを倒したプレイヤーが勝利となる。将棋やチェスでは相手の駒の上に重なるように移動することで相手の駒を取れるが、TUBSTAP では近接して攻撃するか、移動せずに遠隔から攻撃することで相手のユニットを倒す。攻撃対象の相手ユニットを 1 回の攻撃で必ず倒せるわけではない。攻撃を受けたユニットはユニット毎に設定された体力 (HP) が減少し、HP が 0 になって初めてそのユニットが倒れたものとなる。また、近接攻撃を受けて HP が残ったユニットはそのターンで反撃を行う。

TUBSTAP では盤のサイズやユニットの初期配置位置が 1 つに決まっておらず、プレイヤーが任意に決めたマップでゲームを開始する。盤上の各マスには道路・平原・陣地・森・山・海などの地形属性があり、これらは移動に制約を加えたり、攻撃の戦果に影響を与える。

### 2.1.1 ユニットの種類

ゲーム内で用いるユニットは 6 種類あり、それぞれユニット間の相性や移動できる範囲が異なる。

- 戦闘機 (F)：航空ユニット。対空ユニット。A に強く、R に弱い。
- 攻撃機 (A)：航空ユニット。対地ユニット。P,U に強く、F,R に弱い。
- 戦車 (P)：地上ユニット。対地ユニット。R,U に強く、A に弱い。
- 自走砲 (U)：地上ユニット。遠距離攻撃ユニット。範囲内の地上ユニットを一方向的に攻撃できる
- 対空戦車 (R)：地上ユニット。F,A に強い対空ユニットだが、対地攻撃も可能。
- 歩兵 (I)：地上ユニット。歩兵以外のへの攻撃力は期待できないが、壁になる。

## 2.2 戦闘の詳細

自分ユニットが相手ユニットに攻撃を行った際の戦闘結果は、ユニット間の相性、ユニットの HP、地形効果で決定される。表 1 はユニット間の相性を表す。

表 1: ユニットの相性表

ユニット	射程	攻撃力					
		F	A	P	U	R	I
戦闘機 F	1	55	65	0	0	0	0
攻撃機 A	1	0	0	105	105	85	115
戦車 P	1	0	0	55	70	75	75
自走砲 U	2-3	0	0	60	75	65	90
対空戦車 R	1	70	70	15	50	45	115
歩兵 I	1	0	0	5	10	3	55

地上ユニット (P,U,R,I) は、平原で 1、森で 3、山と陣地で 4 で地形効果を受ける。航空ユニットは地形効果を受けない。これらの要素から、最終的な戦闘ダメージは以下の式で定まる。

$$\text{ダメージ} = \frac{(\text{攻撃力} \times \text{攻撃ユニット HP} + 70)}{(100 + (\text{地形効果} \times \text{防御ユニット HP}))}$$

## 3 既存のア戦略ルゴリズム

### 3.1 最良単独行動

TUBSTAP の標準搭載 AI で用いられているアルゴリズム [1]。決められたユニット毎に評価値が最も高い攻撃行動を順に選択し、攻撃行動がなければ相性の良い敵ユニットに近づくか、有利な地形に留まるように移動行動をとる。

### 3.2 行動ペアの合計評価

村山らは最良単独行動を改良した行動ペアの合計評価にもとづく選択を行なった [6]。順にユニットを 2 つ選択し、二つの行動の結果、そのときの評価値の合計が最大になるように 1 番目のユニット行動を選択決定する。その後、未行動のユニットグループについて同様の手順を繰り返す。

### 3.3 深さ限定モンテカルロ法

深さ限定モンテカルロ法は、藤木らによって提案されたアルゴリズム [2] である。すべてのユニットの行動の組み合わせ全体の中からランダムサンプリングし、そのターンの行動を決め、次の相手ターンの行動以降をモンテカルロシミュレーション (プレイアウト) する。シミュレーションは限定した深さで途中で打ち切り、生存ユニットの HP から盤面の優劣を評価する。ランダムサ

ンプリングしたものの中で、評価値が最も高い行動の組み合わせを選択する。

深さ限定モンテカルロ法

1. 盤面の状態  $s$  における全てのユニットの行動の組み合わせ  $a$  の集合  $A(s)$  からランダムに  $m$  個サンプルし、その集合を  $A'(s)$  とする。

$$A'(s) = a_1, a_2, \dots, a_m \subset A(s) \quad (1)$$

2.  $A'(s)$  に含まれる  $a$  による次局面  $s'(s, a)$  を  $d$  ターン先までシミュレーションし、状態評価関数  $g(s)$  により評価を行う。これを  $n$  回繰り返す、評価値の合計を測定する。

$h_i(s) : S \rightarrow S$   $i$  回目のシミュレーションを  $d$  ターン進める。 $S$  は全ての盤面の状態  $s$  の集合。

$g(s) : S \rightarrow R$  評価する。

$$g_n(a) = \sum_1^n g(h_i(s'(s, a)))$$

3.  $A'(s)$  の中で評価最大の行動  $a^*$  を選択する。

$$a^* = \operatorname{argmax}_{a \in A'(s)} g_n(a) \quad (2)$$

#### 4 ユニット行動木 UCT 探索

全体での探索は広すぎるため、ユニット順を考慮しつつ探索するユニット行動木 UCT 探索を提案する。ユニット行動木 UCT 探索とは、ユニット単位で展開する探索木を生成し、その探索に UCT 探索 [7] を用いたものである。UCT 探索は、ゲーム木のノード評価にモンテカルロシミュレーションを用い、有望なノードを逐次的に展開することで、より精度の高い確率的探索を行う手法である。とくに囲碁の人工プレイヤーで高い効果を上げている。

##### 4.1 ユニット行動木

ユニット行動木は、1つのノードに1つユニットの行動を割り当てた探索木である。未行動ユニットが複数あるとき、図2に示すような自分のターンが複数段に展開された探索木となる。図において点線までの木は自分の1ターンである。

ターン制戦略ゲームでは、1ターンの行動が複数のユニットの行動の組み合わせ全体であるためその数は膨大である。モンテカルロシミュレーションを行うにしても、すべてのノードについては不可能である。このため、まったくノード展開後の状況を考慮しないで枝刈りされるため、有望な手を見逃す危険性が極めて高い。

ユニット行動木では、1つのノードの枝分かれば、た

だかユニットの行動のユニット数分の総和であり、モンテカルロ探索を行えば、各行動の評価推定値を得る事ができる。その上で、ノード展開や枝刈りを行うことが可能となる。

#### 4.2 ユニット行動木 UCT 探索のアルゴリズム

##### ユニット行動木 UCT 探索

1. 局面の状態  $s$  を根ノードの盤面の状態とする
2.  $s$  における未行動のユニットの行動の集合  $A(s)$  に含まれる  $a$  による次局面  $s'(s, a)$  が、 $s$  の子ノードとなる
3.  $A(s)$  から UCB 値が最大の行動  $a^*$  の枝をたどり、子ノードに向かう

$$UCB = X_i + c \sqrt{\frac{\log N}{n_i}} \quad (3)$$

$$a^* = \operatorname{argmax}_{a \in A(s)} UCB(a) \quad (4)$$

4. 選択された子ノードが末端でなければ、 $s = s'(s, a^*)$  として2に戻る
5. たどり着いた末端ノードのプレイアウト回数が閾値を超えていた場合、1段展開し4と同様に2に戻る
6. プレイアウト回数が閾値を超えていない末端ノードにたどり着いたらプレイアウトを行う
7. プレイアウト結果をたどってきたノードに反映する

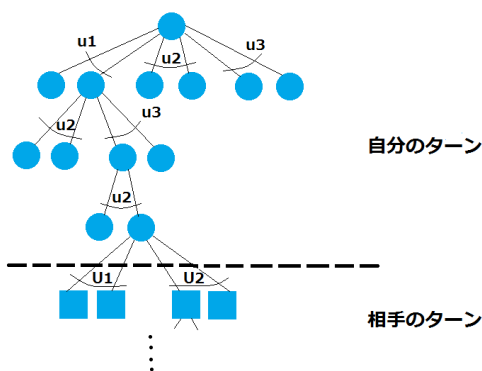


図2: ユニット行動木 UCT 探索の探索木

#### 5 提案アルゴリズムの評価

対戦実験の勝率から、提案手法が既存の手法の深さ限定モンテカルロ法 [2] に対し優位であるかを確認する。実験盤面は、TUBSTAP 付属のものから様々な広さ

の以下の3つのマップを選び、それぞれマップ a, マップ b, マップ c とした。

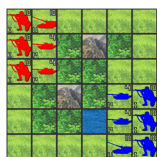


図 3: マップ a

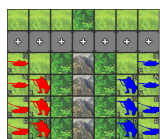


図 4: マップ b



図 5: マップ c

### 5.1 実験設定

対戦回数は 100 回。ターン上限数はマップ a は 30, マップ b, マップ c は 20 とし, 上限数を越えた時の残りユニットの合計 HP が 10 以上なら合計 HP が多いチームの勝利とし, それ以外は引き分けとした。

深さ限定モンテカルロ法は, 行動の組み合わせのサンプル数  $m$  は 100, シミュレーションの深さ  $d$  は 3 ターン先まで行い, 各サンプル毎にモンテカルロシミュレーションを 80 回行った。ユニット行動木 UCT 探索は, 総プレイアウト数を 2000, 各ノードのプレイアウト回数の閾値を 20, UCB 値の式の定数  $c$  を 0.15 とした。

### 5.2 実験結果

対戦結果を表 2 に示す。勝率の計算は引き分けを 0.5 として計算した。

表 2: ユニット行動木 UCT 探索 vs 深さ限定モンテカルロ法 (各マップ 100 試合)

	勝	負	引分	勝率
マップ a	51	40	9	0.56
マップ b	58	13	29	0.73
マップ c	73	8	19	0.83

### 5.3 考察

すべてのマップにおいて提案手法のユニット行動木 UCT 探索が勝ち越しており, また, マップ a からマップ c にかけて勝率が増加していることが分かる。盤面の大きさやユニットの種類が増加するほど合法手の数も増加し, その分時間あたりに探索できる木の深さは浅くなる。それにも関わらず勝率が増加しているということは, プレイアウトの結果を保証するのに求められる探索木の深さはかなり浅く, 浅い手を広く万編に探索することが効果的であるということが考えられる。また, プレイアウトを終局まで行ったことも勝率を上げた原因のひとつであると考えられる。

## 6 ゲーム局面の大局的ファジィ評価

分岐因子が大きく探索が困難な戦略ゲームに対しては, モンテカルロ探索の有効性も限定的である。その一方で人間プレイヤーはそれなりに強く, コンピュータプレイヤーに対しても勝ち越す現状である。このとき, 人間プレイヤーがコンピュータより深く広い探索を行っていることはほぼあり得ず, 直感的な局面判断に優れ, 大幅な探索の枝刈りができているためと考えられる。

この, 人が直感的に行うゲーム局面の大局的な評価をファジィ集合を介して行うファジィ評価を提案する。

ファジィ集合は人のもつ概念をある要素がどれだけその集合に含まれるかという帰属度 (メンバーシップ値) によって特徴付けて, あいまいな集合として定式化したものである。ゲーム局面の評価においては, 人の知識にもとづいて良い盤面というあいまいな集合を定義し, ゲーム中, 新たに得られた局面の良い盤面集合への帰属度を参照することで人の直感にもとづいたその局面の良さを評価することができる。

### 6.1 ファジィ集合

ファジィ集合とはファジィ理論の根幹をなす概念で, 大きな数のような境界のあいまいな集合である。Zadeh はこのようなファジィ集合  $A$  を全体集合  $X$  の上で, 式 (5) に示すメンバーシップ関数  $\mu_A$  を定義することで特徴付けた。

$$\mu_A : X \rightarrow [0, 1] \quad (5)$$

たとえばある要素  $x \in X$  があいまいな集合  $A$  に含まれる度合いは  $\mu_A(x) = 0.8$  などとして表現される。

### 6.2 ファジィ集合の同定と生成

ファジィ集合を構築することはメンバーシップ関数を定義することと同値である。メンバーシップ関数の同定にはいくつかの方法が提案されており, たとえば以下のようなものがある。

- 関数形を人が試行錯誤を経て直接生成する
- ファジィ数を基礎として生成する
- データに基づき生成する

ゲーム局面の状態空間は多変数で離散的なため, 人手によったり, 一次元のファジィ数をもとにすることは困難である。ある局面がどのように分類されるかについては, 人の直感的な判断を得る事ができるので, データに基づいたファジィ集合の生成を行う。

たとえば特徴量ベクトル  $v_i$  の空間  $V$  上で, 良い局面というファジィ集合  $G$  を生成することを考える。

以下の手順で生成することができる。

#### ファジィ集合の生成

1. 局面  $s_i$  を一つ生成する
2. 人間プレイヤーが  $G$  に含まれると言えるかどうか判断し  $h(s_i) \in \{0, 1\}$  として含まれるかどうかを記録する
3. 十分なサンプル集合  $S = \{s_1, s_2, \dots, s_n\}$  が集まるまで 1 へ戻る
4. 得られたサンプル  $s_i$  を特徴量ベクトル  $\mathbf{v}_i = \mathbf{v}(s_i)$  に変換する
5.  $S \rightarrow V$  は多対 1 の写像となるため、同一の特徴量ベクトル  $\mathbf{v}_k$  をあたえるサンプル集合  $S_k = \{s_j | V(s_j) = \mathbf{v}_k\}$  を集計し、 $\mathbf{v}_k$  の生成目標の集合  $G$  への帰属度を式 (6) の比率によって与える
6. すべての  $\mathbf{v}_i \in V$  についての帰属度を表  $T_G$  として保持する

$$\mu_G(\mathbf{v}_k) = \frac{|\{s_j | h(s_j) = 1, \forall s_j \in S_k\}|}{|S_k|} \quad (6)$$

実装上は、ここで生成された帰属度の表  $T_G$  がファジィ集合  $G$  のメンバーシップ関数の実体である。表に含まれない一般の特徴量ベクトル  $v \in V$  については値の定まった他の要素から補完によって求める。

## 7 ファジィ評価による枝刈り

ユニット行動木 UCT 探索にファジィ判断の評価による候補手の枝刈りを加えた。枝刈りを行わないユニット行動木 UCT 探索と、その効果を実験的に比較分析する。実験盤面は、5 節で用いたマップ a を選んだ。

### 7.1 ファジィ評価

人間の直感的に選択する行動を AI も選択できるように、行動のファジィ評価を行う。まず、行動を以下の特徴量ベクトル  $\mathbf{v}$  で表す。

- $v_0$  : 行動したユニットの種類
- $v_0 \in \{1, 2, 3, 4, 5, 6\}$
- $v_1$  : 行動したユニットの HP
- $1 \leq v_1 \leq 10$
- $v_2$  : 攻撃対象ユニットの種類
- $v_2 \in \{0, 1, 2, 3, 4, 5, 6\}$
- $v_3$  : 攻撃ダメージ
- $0 \leq v_3 \leq 10$
- $v_4$  : 次のターンに受ける合計ダメージ

$$0 \leq v_4 \leq 10$$

次に、人間プレイヤーと AI の対戦を行う。行動の主観評価を収集するため、対戦内で、人間プレイヤーが選択した行動は正例として保存する。さらに AI が選択した行動の中から、対戦相手の人間プレイヤーが自分だったらそのように選択すると考える行動を正例とし、明らかに悪手であると断定できる行動は捨て、そのどちらとも言えないよく分からないと思った行動には重み 0.5 の正例とする。

人間プレイヤーを教師としているため、人間プレイヤーは対戦相手の AI よりも十分に強いことが求められる。予備実験として筆者とユニット行動木 UCT 探索の対戦成績を表 3 に示す。この結果から、人間の方が強いと言える。

表 3: 人間プレイヤー (筆者) vs AI (ユニット行動木 UCT) (10 試合)

	勝	負	引分	勝率
人間プレイヤー	3	1	6	0.6

引き分けの割合が多いのは、筆者の戦略が持久戦のためである。

対戦を繰り返し行い行動データを収集し、特徴量ベクトル空間において良い行動というファジィ部分集合を生成する。今回は補完にもちいる周辺データとして  $k$ -近傍を 0 として、同一ベクトルでの正例密度を計算した。サンプル数は 440 である。収集したデータの中に含まれない種類の行動は、特徴量がユークリッド距離で最も近い行動のファジィメンバーシップ値を補間し、ファジィ評価値とした。こうして全ての行動にファジィ評価値を定め、ある閾値を下回る行動を探索木から枝刈りし、より深い探索を行う。本実験で求めたメンバーシップ値の表の一部を表 4 に示す。

表 4: 行動のメンバーシップ値の例

	$v_0$	$v_1$	$v_2$	$v_3$	$v_4$	$\mu_G(v)$
$\mathbf{v}_1$	2	3	2	2	3	0.71
$\mathbf{v}_2$	2	4	2	2	4	0.8
$\mathbf{v}_3$	3	10	-1	0	0	1
$\mathbf{v}_4$	5	2	5	2	2	0.33
$\mathbf{v}_5$	5	10	2	1	10	0
$\vdots$			$\vdots$			$\vdots$

## 7.2 実験設定

枝刈り有り無しでの直接対決実験を行った。対戦回数は500回。ターン上限数は20とし、上限数を越えた時の残りユニットの合計HPが10以上なら合計HPが多いチームの勝利とし、それ以外は引き分けとした。マップが非対称であるため、先攻後攻を250回ずつである。

両ユニット行動木UCT探索は、1行動あたりのプレイアウト数は2000、各ノードのプレイアウト回数の閾値は10、UCB値の式の定数 $c$ は0.15、枝刈りの閾値は0.75とし、違いは枝刈りの有無のみである。

## 7.3 実験結果

対戦結果を表5に示す。勝率の計算は引き分けを0.5として計算した。

表 5: ユニット行動木 UCT 探索 (枝刈り有) vs ユニット行動木 UCT 探索 (枝刈り無) (500 試合)

	勝	負	引分	勝率
枝刈り有	189	235	76	0.454

## 7.4 考察

実験結果から、枝刈り有りの勝率が5割を下回っており、枝刈りを行うことで強くはならなかった。

枝刈りにより、最初のターンの探索木の深さは表6のように変化した。

表 6: 最初のターンの探索木の深さ

	枝刈り無	枝刈り有
1 手目	2.03	6.15
2 手目	2.21	5.30
3 手目	3.30	4.94
4 手目	3.00	5.00
5 手目	2.63	4.97
6 手目	2.33	4.15

枝刈りを行うことで2~4段深く探索を行えていることが分かった。

対戦実験中の両AIが取る行動の観察から、枝刈り有りの方は枝刈り無しの方に比べて、人間プレイヤーが選択しないような突拍子の無い行動が少なく、概ね人間が予測する行動を選択していることが分かった。

枝刈りの有無に関わらずプレイアウト数を2000に固定しているが、今回の実験設定ではプレイアウトの時間に比べて枝刈りの時間ははるかに短いからである。別の実験設定で、探索木の深さが増して枝刈りの回数が増加すれば、計算時間に影響されることが予想される。

## 7.5 ファジィ評価値のランダムシャッフルとの比較

ファジィ評価値の値の有効性を評価するため、それぞれの行動のファジィ評価値をランダムシャッフルし、再び割り当てたもので枝刈りを行ったもので対戦実験を行った。

実験設定は7.2節と同じとし、表7に対戦結果を示す。

表 7: ユニット行動木 UCT 探索 (ランダム枝刈り) vs ユニット行動木 UCT 探索 (枝刈り無) (500 試合)

	勝	負	引分	勝率
ランダム枝刈り	185	237	78	0.448

ファジィ評価値をランダムシャッフルしても勝率に大きな変化は現れなかった。このことから、本来のファジィ評価値に大局観を有効に表しているとは結論づけられない。

次に、ファジィ評価値のシャッフルを行ったものとシャッフルを行わずそのまま用いたもので対戦実験を行った。

学習した行動データ数を1246に増加し、その他の実験設定は上記実験と同じものとした。表8が対戦結果である。

表 8: ユニット行動木 UCT 探索 (枝刈り) vs ユニット行動木 UCT 探索 (ランダム枝刈り) (500 試合)

	勝	負	引分	勝率
枝刈り	213	205	82	0.508

実験結果から、互角の強さであることが分かる。ファジィ評価値をそのまま用いたものがシャッフルして用いたものと互角の強さであるということからも、今回のファジィ評価値は大局観を表せていないことが考えられる。

## 8 ファジィ評価値のUCB値への適用

次に、ファジィ評価値をUCBの勝敗初期値として用いて対戦実験を行う。親ノードから展開されて、一度も探索が行われていない子ノードのUCB初期値をファジィ評価値とし、一度でも探索を行ったノードは式(3)に従って定める。

$$UCB = \begin{cases} \mu_G(v) \in [0, 1] & n_i = 0 \\ X_i + c\sqrt{\frac{\log N}{n_i}} & o.w. \end{cases} \quad (7)$$

こうして未探索のノードの中からファジィ評価値の高いものを優先して探索を行うようにする。実験設定は、

対戦回数 200 回，その他は枝刈りに用いた場合と同じである．

### 8.1 実験結果

対戦結果を表 9 に示す．勝率の計算は引き分けを 0.5 として計算した．

表 9: ユニット行動木 UCT 探索 (UCB 値ファジィ) vs UCB 値そのまま

	勝	負	引分	勝率
UCB 値にファジィ評価値	65	98	37	0.4175

### 8.2 考察

UCB 値にファジィ評価値を影響させても，勝率は上昇しなかったが，探索深さは最大 40 超であった．

## 9 おわりに

本研究ではターン制戦略ゲームの研究用に提案された TUBSTAP を用い，ユニット行動木 UCT 探索及びその枝刈りにおけるファジィ評価の適用を検討した．ユニット行動木 UCT 探索は，深さ限定モンテカルロ法より強くなることを実験的に確かめた．人の大局観をファジィ集合を用いて表すことでファジィ評価による枝刈りを提案した．実験では探索は深くなったものの，勝率には結びついていない結果が得られた．しかし，AI の行動着手の挙動に変化は見られ，まだ改善の余地があると考えられる．今後は，ファジィ評価をより多くのデータから求めて精度を高めるとともに，補間方法を検討することが課題である．[4]

本研究の一部は科学研究費補助金基盤研究 C(26330273) により支援されたものである．

## 参考文献

- [1] ターン制戦略ゲーム 学術用基板プロジェクト tubstap. <http://www.jaist.ac.jp/is/labs/ikedalab/tbs/>.
- [2] 翼藤木, 公志朗村山, 心池田. ターン制ストラテジーにおける状態評価関数を用いた深さ限定モンテカルロの適用. エンターテイメントと認知科学研究ステーション (E&C) 第 8 回シンポジウム, pp. -, 2014.
- [3] 加藤千裕, 三輪誠, 鶴岡慶雅, 近山隆. ターン制ストラテジーゲームにおける戦術決定のための uct 探索とその効率化. 第 18 回ゲームプログラミングワークショップ, pp. 138–145, 2013.

- [4] Lotfi A Zadeh. Fuzzy sets. *Information and control*, Vol. 8, No. 3, pp. 338–353, 1965.
- [5] Lotfi A Zadeh. The concept of a linguistic variable and its application to approximate reasoning i. *Information sciences*, Vol. 8, No. 3, pp. 199–249, 1975.
- [6] 村山公志朗, 藤木翼, 池田心. 学術研究用プラットフォームとしての大戦略系ゲームのルール提案. 第 18 回ゲームプログラミングワークショップ, pp. 146–153, 2013.
- [7] C. B. Browne, E. Powley, D. Whitehouse, S. M. Lucas, P. I. Cowling, P. Rohlfshagen, S. Tavener, D. Perez, S. Samothrakis, and S. Colton. A survey of monte carlo tree search methods. *Computational Intelligence and AI in Games, IEEE Transactions on*, Vol. 4, No. 1, pp. 1–43, march 2012.